

UNIVERSIDADE DE LISBOA

Faculdade de Medicina



**O papel da LARP4 no Sistema nervoso central: relação com a
Atrofia Muscular Espinal**

João Paulo de Sá e Silva

Orientadores: Professora Doutora Margarida Gama-Carvalho

Professora Doutora Ana Sebastião

Dissertação especialmente elaborada para obtenção do grau de Mestre em
Neurociências

2018

UNIVERSIDADE DE LISBOA

Faculdade de Medicina



**O papel da LARP4 no Sistema nervoso central: relação com a
Atrofia Muscular Espinal**

João Paulo de Sá e Silva

Orientadores: Professora Doutora Margarida Gama-Carvalho

Professora Doutora Ana Sebastião

Dissertação especialmente elaborada para obtenção do grau de Mestre em
Neurociências

2018

Todas as afirmações efetuadas no presente documento são da exclusiva responsabilidade do seu autor, não cabendo qualquer responsabilidade à Faculdade de Medicina de Lisboa pelos conteúdos nele apresentados.

“A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de 18 de Setembro de 2018.”

*"However bad life may seem, there is
always something you can do and succeed at.
Where there's life, there's hope."*

Stephen Hawking

*"Take the first step in faith. You don't have
to see the whole staircase, just take the first step."*

Martin Luther King Jr.

Os resultados da presente dissertação fazem parte das seguintes comunicações sob a forma de poster:

João Sá, Mariana Oliveira, Ana Luísa Gomes, Ana Margarida Matos, Isabel Peixeiro, Gonçalo Costa, Carlos Cordeiro, Aaron Voigt and Margarida Gama-Carvalho (2017) Spinal Muscular Atrophy: is LARP4 a positive modifier of SMN expression?. **21^a Reunião Anual Sociedade Portuguesa Genética Humana**. Capuchos, Almada, Portugal, Nov 16-18, 2017;

João Sá, Mariana Oliveira, Ana Luísa Gomes, Ana Margarida Matos, Isabel Peixeiro, Gonçalo Costa, Carlos Cordeiro and Margarida Gama-Carvalho (2017) Spinal Muscular Atrophy: LARP4 is a novel regulator of SMN expression. **XI Meeting of the Portuguese Society for Neuroscience**. Braga, Portugal, Maio 25-26, 2017;

João Sá, Mariana Oliveira, Ana Luisa Gomes, Isabel Peixeiro and Margarida Gama-Carvalho (2016). A role of LARP4 in the regulation of SMN expression. **UCL Neuroscience 2016**. Londres, Reino Unido, Junho 24, 2016.

Agradecimentos

Chegado o término deste grande percurso, quero agradecer desta forma às pessoas que confiaram em mim e me apoiaram neste tempo de trabalho bastante árduo.

Primeiramente gostava de agradecer à minha orientadora, Professora Doutora Margarida Gama-Carvalho, por me ter aceite no seu grupo de investigação e por ter confiado em mim e na minha vontade de aprender a fazer ciência. Agradeço profundamente do fundo do meu coração.

De seguida, queria agradecer à minha co-orientadora, Professora Doutora Ana Sebastião, pela sua disponibilidade e por me ter aceite como seu estudante e por me ter ajudado e por ter sido minha “madrinha” na Sociedade Portuguesa de Neurociências.

I would like to thank Aaron Voigt and his team for the availability and help in developing the work with the animal model *Drosophila melanogaster* and in the opinions and suggestions, thank you very much.

Também gostaria de agradecer à Professora Doutora Solveig Thorsteinsdóttir e sua equipa de investigação pelo acolhimento fantástico e na colaboração e ajuda na execução dos ensaios de imunohistoquímica, o meu muito obrigado.

O meu percurso em investigação começou perto da minha terra natal, em Braga, no Instituto de Investigação em Ciências da Vida e da Saúde da Universidade do Minho onde agradeço à Professora Joana Palha por me ter acolhido no Programa Doutoral. O meu muito obrigado ao Doutro Bruno Costa e ao Doutor Tiago Oliveira, e vossas equipas, pelo acolhimento fantástico e pela aprendizagem que fui desenvolvendo ao longo do meu percurso no ICVS, assim como pela vontade de desenvolver ciência em Neurociências.

Em Lisboa, conheci pessoas fantásticas que estiveram, e continuam a meu lado, e me apoiaram nos momentos bons e nos momentos menos bons durante estes anos. Agradeço a todos vós pela paciência com que me aturaram e peço desculpa pela minha ausência, por vezes marcada; a todos vocês que estão no meu coração e vão ficar para sempre: Teresa D’Oliveira, Ramadas (Lourenço Maria e Duarte), Vânia, Inês (Xu), Inês Nabais, Sara, Baleizão, Condessa, Manu, Diniz, Rita Mendes, Paulo Gonçalves, Pedro’s, Tiago, João Pedro Santos, Inês Garcia, Nuno, Medeiros,

Cris (Brasileira), Ângela, Catarina, Cláudia, Geina, Dr^a Isabel, Sr. Paulo e toda a sua família (Ana, Hugo e D. Luísa), o meu muito obrigado.

Para além das pessoas mencionadas anteriormente, há um conjunto de outras pessoas fantásticas, que conheci durante o meu percurso no laboratório de acolhimento (Verónica, Sofia, Celma, Inês, Zé, Luís Marques), e que sem a ajuda e o apoio delas o meu percurso não teria sido possível como foi e a elas agradeço imenso o tempo que disponibilizaram a ouvir os meus desabafos e as minhas frustrações deste meu percurso árduo, o meu muito obrigado.

Também gostava de agradecer a uma pessoa que me ajudou nos momentos de maior ansiedade, stress e frustração e meu deu o incentivo e a vontade da prática do desporto e a quem me sinto realmente grato por tudo o que fez. Muito obrigado João Teixeira por todos aqueles momentos que eu pedia mais tempo para descansar e tu “obrigavas-me” a fazer mais uma série de repetições.

Não poderia de deixar de agradecer a três pessoas muito especiais na minha vida que conheci e que estão e estarão sempre no meu coração. Queria agradecer por tudo o que fizeram por mim e pedir desculpa por alguma ausência. Agradeço-te Ramadas, Duarte e ao meu querido Lourenço Maria a felicidade que me fornecem quando estou convosco.

E para último deixo profundamente a minha gratidão a um núcleo de pessoas que, infelizmente, não estão próximas fisicamente, mas estão sempre no meu coração e sempre me apoiaram nas minhas decisões e sempre preocupadas comigo e com o desenvolvimento do meu trabalho. O meu muito obrigado à minha Mãe (Mi), ao meu pai (Zinho), ao meu irmão André, à minha tia Clara que a adoro e às minhas primas Nicoles e Becas que sempre me apoiaram no meu percurso. A vocês agradeço do fundo do meu coração e estão sempre comigo.

Por fim, queria dedicar esta minha dissertação a uma pessoa que infelizmente não se encontra entre nós, mas que gosto imenso e tenho imensas saudades. Um beijinho grande de muitas saudades à minha querida avó – Maria Helena.

ÍNDICE

| | |
|--|-----------|
| AGRADECIMENTOS..... | 11 |
| ABSTRACT | 15 |
| RESUMO | 17 |
| ÍNDICE DE FIGURAS..... | 19 |
| ÍNDICE DE TABELAS..... | 23 |
| LISTA DE ABREVIATURAS..... | 25 |
| I. INTRODUÇÃO..... | 27 |
| 1.1 NEURÓNIO MOTOR: DA NEUROBIOLOGIA À CLÍNICA | 29 |
| 1.1.1 <i>Biologia e Histologia do Neurónio Motor</i> | 29 |
| 1.1.2 <i>As Doenças do Neurónio Motor</i> | 32 |
| 1.1.2.1 <i>Epidemiologia das doenças do neurónio motor</i> | 32 |
| 1.1.2.2 <i>Classificação e clínica das doenças do neurónio motor</i> | 33 |
| 1.1.2.3 <i>Mecanismos fisiopatológicos das doenças do Neurónio Motor</i> | 34 |
| 1.2 ATROFIA MUSCULAR ESPINAL | 38 |
| 1.2.1 <i>Atrofia Muscular Espinal e sua clínica</i> | 38 |
| 1.2.1.1 <i>Manifestações Clínicas</i> | 39 |
| 1.2.1.2 <i>Diagnóstico e meios complementares de diagnóstico</i> | 41 |
| 1.2.2 <i>Bases moleculares da Atrofia Muscular Espinal</i> | 42 |
| 1.2.2.1 <i>Modificador da severidade da SMA: SMN2</i> | 44 |
| 1.2.2.2 <i>Proteína SMN e suas funções</i> | 46 |
| 1.2.2.3 <i>Proteína SMN e o Sistema nervoso central</i> | 48 |
| 1.2.3 <i>Modelos Animais</i> | 49 |
| 1.3 PROTEÍNAS DE LIGAÇÃO AO RNA (<i>RNA BINDING PROTEINS/RBP</i>) E AS DOENÇAS NEURODEGENERATIVAS | 52 |
| 1.3.1 <i>Regulação da expressão génica no processo de transcrição e de processamento do RNA mensageiro</i> | 53 |
| 1.3.2 <i>Regulação da expressão génica no processo de tradução</i> | 58 |
| 1.3.3 <i>Proteínas de ligação ao RNA (RNA Binding Proteins)</i> | 61 |
| 1.3.3.1 <i>Proteína LARP4 e sua família</i> | 64 |
| 1.4 <i>DROSOPHILA MELANOGASTER</i> COMO MODELO ANIMAL DAS DOENÇAS DO NEURÓNIO MOTOR | 66 |
| 1.4.1 <i>Ciclo de vida</i> | 68 |
| 1.4.2 <i>O sistema nervoso central da Drosophila melanogaster</i> | 69 |
| 1.4.3 <i>Genoma e ferramentas de manipulação genética</i> | 71 |
| 1.4.3.1 <i>GAL4/UAS como um sistema de expressão génica binário</i> | 73 |
| II. OBJETIVOS..... | 81 |
| III. MATERIAIS E MÉTODOS..... | 85 |
| 3.1 ESTUDO DA EXPRESSÃO E FUNÇÃO DA LARP4..... | 87 |
| 3.1.1 <i>Linhas celulares e cultura de células</i> | 87 |
| 3.1.2 <i>Transfecção celular</i> | 87 |
| 3.1.3 <i>Ensaio de Luciferase</i> | 88 |
| 3.1.4 <i>Isolamento e extração de RNA</i> | 89 |

| | | |
|--------|--|------------|
| 3.1.5 | <i>Transcrição reversa e PCR em tempo real</i> | 89 |
| 3.1.6 | <i>Extratos proteicos</i> | 90 |
| 3.1.7 | <i>Western-Blot</i> | 90 |
| 3.1.8 | <i>Imunocitoquímica</i> | 91 |
| 3.1.9 | <i>Imunohistoquímica</i> | 92 |
| 3.1.10 | <i>Análise Estatística</i> | 92 |
| 3.2 | ESTUDO DA FUNÇÃO DA PROTEÍNA CG11505 | 93 |
| 3.2.1 | <i>Restrição enzimática, vetor e insert</i> | 93 |
| 3.2.2 | <i>Reação de ligação</i> | 94 |
| 3.2.3 | <i>Transformação bacteriana</i> | 94 |
| 3.2.4 | <i>Extração e purificação de plasmídeos (Miniprep)</i> | 94 |
| 3.2.5 | <i>Eletroforese em gel de agarose</i> | 95 |
| 3.2.6 | <i>Sequenciação</i> | 95 |
| 3.2.7 | <i>Análise bioinformática</i> | 96 |
| 3.2.8 | <i>Linhas de sobre-expressão e knock-down</i> | 96 |
| 3.2.9 | <i>Análise funcional da proteína CG11505 em Drosophila</i> | 97 |
| IV. | RESULTADOS | 99 |
| 4.1 | ANÁLISE BIOINFORMÁTICA DA HOMOLOGIA ENTRE CG11505 E LARP4 | 101 |
| 4.2 | LOCALIZAÇÃO CELULAR DA PROTEÍNA LARP4 | 103 |
| 4.3 | ANÁLISE HISTOLÓGICA DA LOCALIZAÇÃO DA PROTEÍNA LARP4 NAS JUNÇÕES NEUROMUSCULARES EM EMBRIÕES DE RATINHO | 106 |
| 4.4 | ANÁLISE DO PAPEL DA LARP4 NA REGULAÇÃO DA EXPRESSÃO DA PROTEÍNA SMN EM MODELOS CELULARES..... | 109 |
| 4.4.1 | Efeito da expressão da LARP4 sob o gene repórter Luciferase- 3'UTR do SMN1/2..... | 110 |
| 4.4.2 | Efeito da expressão da LARP4 sob o SMN endógeno | 112 |
| 4.5 | ESTUDO DA FUNÇÃO DO HOMÓLOGO DE DROSOPHILA DA PROTEÍNA LARP4 NO SISTEMA NERVOSO | 116 |
| 4.5.1 | <i>Clonagem do vetor de expressão pUAST_CG11505</i> | 116 |
| 4.5.2 | <i>Análise do impacto da sobre-expressão do gene CG11505 em Drosophila melanogaster</i> | 118 |
| 4.5.3 | <i>Análise do impacto da sub-expressão do gene CG11505 em Drosophila melanogaster</i> | 122 |
| V. | DISCUSSÃO | 123 |
| VI. | CONCLUSÃO E PERSPETIVAS FUTURAS | 131 |
| VII. | REFERÊNCIAS BIBLIOGRÁFICAS | 135 |
| VIII. | ANEXO | 149 |
| | ANEXO I: SEQUÊNCIA DO VETOR DE EXPRESSÃO PUA _{ST} _CG11505..... | 151 |

Abstract

Spinal Muscular Atrophy (SMA) is a severe neurodegenerative disease characterized by the degeneration of motor neurons located in the anterior horn of the spinal cord due to the deletion or mutation of the SMN1 gene and the decrease in SMN protein production. Its incidence is about 1:10,000 live births. LARP4 protein was identified as an RBP that binds to the 3'UTR region of the SMN1/2 gene. Its homolog of *Drosophila melanogaster* - CG11505 - was identified as a positive modifier of SMA phenotype in a genomic screen in a disease model. LARP4 protein shows a cytoplasmic distribution in the HEK293 and SK-N-BE(2)c cell lines; but in the SMA patient-derived fibroblast GM03813 cell line it shows a cytoplasmic and nuclear pattern, in a very close location with the Cajal Bodies. Functional studies with over-expression and knock-down of LARP4 have been performed, and increased expression of LARP4 protein has been shown to be related to increased expression of the endogenous Smn protein in the SK-N-BE(2)c cell line, while its decrease results in a decrease in SMN mRNA levels in HEK293 and SK-N-BE(2)c cell lines. The overexpression result was also observed in HEK293 cell line transfected with a vector containing the 3'UTR region of the SMN1/2 mRNA gene. In this way, it was possible to infer that the increase of LARP4 expression has a role in the regulation of the translation of the Smn protein and that the decrease of its expression leads to a decrease in mRNA levels, possibly due to a decrease in mRNA stability and degradation. In addition, the function of the CG11505 protein in the nervous system was analyzed in *Drosophila melanogaster*. We observed that the decrease of CG11505 protein expression results in the degeneration of retinal neurons. In this way, we propose that the LARP4 protein has a role in the regulation of SMN gene expression, potentially playing a relevant role in the specific context of the nervous system.

Keywords: La-related Protein 4, Spinal Muscular Atrophy, Survival of Motor Neuron, RNA binding protein

Resumo

A Atrofia Muscular Espinal (SMA) é uma doença neurodegenerativa severa caracterizada pela degeneração dos neurónios motores, localizados no corno anterior da medula espinal, devido à deleção ou mutação do gene SMN1 e à consequente diminuição da produção da proteína SMN. A sua incidência é de cerca de 1:10,000 nados vivos. A proteína LARP4 foi identificada como uma 'RNA Binding Protein' (RBP) que se liga à região 3'UTR do gene SMN1/2. Adicionalmente, o seu gene homólogo em *Drosophila melanogaster* - CG11505 - foi identificado como modificador positivo do fenótipo neuromuscular num modelo de SMA. A proteína LARP4 apresenta uma distribuição citoplasmática nas linhas celulares HEK293, SK-N-Be(2)c; mas na linha celular GM03813, derivada de fibroblastos de um doente de SMA, apresenta um padrão citoplasmático e nuclear, numa localização muito próxima com os Corpos de Cajal. Foram realizados estudos funcionais de sobre e sub-expressão da LARP4, tendo-se verificado que o aumento da expressão da proteína está relacionada com o aumento dos níveis da proteína Smn endógena na linha celular SK-N-Be(2)c; e que a sua diminuição resulta numa diminuição dos níveis do mRNA do SMN nas linhas celulares HEK293 e SK-N-Be(2)c. O resultado de sobre-expressão foi igualmente observado na linha celular HEK293 transfetada com um vetor que contém a região 3'UTR do mRNA do gene SMN1/2. Desta forma foi possível inferir que o aumento da expressão da LARP4 tem um papel na regulação da tradução da proteína Smn e que a diminuição da sua expressão leva à diminuição da abundância do mRNA, possivelmente por diminuição da sua estabilidade e degradação. Para além disso, foi analisada a função da proteína CG11505 no sistema nervoso em *Drosophila melanogaster*, tendo-se observado que a diminuição da sua expressão resulta na degenerescência dos neurónios da retina. Desta forma, propõe-se que a proteína LARP4 tenha um papel na regulação da expressão do gene SMN, potencialmente assumindo um papel importante no contexto específico do sistema nervoso.

Palavras-Chave: Atrofia Muscular Espinal, *La-related Protein 4*, *Survival of Motor Neuron*, *RNA binding protein*

Índice de figuras

| | |
|---|-----|
| Figura 1 – Localização anatômica do neurônio motor superior e inferior e esquema do circuito motor. Imagem adaptada de (McKean, Ross, Dressler, & Scheurer, 2017)..... | 31 |
| Figura 2 - Sinais associados a neurônios motores superiores (UMN) e inferiores (LMN) no desenvolvimento das Doenças do Neurônio Motor. Imagem adaptada de Dharmadasa et al. (2017)..... | 34 |
| Figura 4 - Diagrama dos genes humanos SMN1 e SMN2 no cromossoma 5. Os doentes com SMA apresentam deleção ou mutação do gene SMN1 que é responsável pela produção da proteína SMN. O gene SMN2, devido a uma transição (c840 C>T) no exão 7, resulta na produção maioritária de uma proteína SMN truncada não funcional, rapidamente degradada. Esquema adaptado de Kolb & Kissel (2011) e SUMNER, PAUSHKIN, & KO (2017)..... | 44 |
| Figura 5 - Genótipo e correlação do número de cópias do gene SMN2 com os tipos de SMA. Imagem adaptada de Baioni & Ambiel (2010)..... | 46 |
| Figura 5 – Estrutura do mRNA de eucarióticas com exemplo de alguns elementos que afetam a expressão génica. Imagem adaptada de Mignone et al. (2002).61 | 61 |
| Figura 7 - Ciclo de vida da <i>Drosophila melanogaster</i> . Figura adaptada de Council (2000)..... | 69 |
| Figura 8 - Esquema da organização dos cromossomas de <i>Drosophila</i> . Imagem adaptada de (Hales et al., 2015) | 72 |
| Figura 9 - Esquema do sistema de expressão binário GAL4/UAS. Imagem adaptada de Tan & Azzam (2017) | 74 |
| Figura 10 - Esquemas de modificações do sistema de expressão binário GAL4/UAS: Sistema TARGET (A) e Sistema GAL-ER e GeneSwitch (B). Imagem adaptada de Tan & Azzam (2017) | 79 |
| Figura 11 - Alinhamento das proteínas CG11505 e LARP4, com recurso ao site flyrnai.org..... | 102 |
| Figura 12 – Imunofluorescência das proteínas SMN, LARP4 e coloração de DNA (DAPI) nas linhas celulares HEK293 (A), SK-N-BE(2)c (B), GM03813 (C) e VA-13 (D). A proteína SMN tem uma distribuição nuclear (seta azul). Na linha celular GM03813 observa-se uma proximidade entre as proteínas LARP4 e SMN no núcleo (seta amarela). | 104 |
| Figura 13 -Imunofluorescência das proteínas Coilin, LARP4 e coloração de DNA (DAPI) na linha celular GM03813. A proteína Coilin tem uma distribuição nuclear, nos Corpos de Cajal (seta roxa) e a LARP4 e a Coilin não se localizam no mesmo local (seta amarela). | 105 |
| Figura 14 - Alinhamento das proteínas LARP4 humana e Larp4 de ratinho, com recurso ao site flyrnai.org..... | 107 |
| Figura 15 – Imunohistoquímica da localização da LARP4 na junção neuromuscular em embrião de ratinho, E15.5, na região pré- (D-E) e pós-sináptica (A-B) com ampliação de 20x e 60x. As imagens C e F representam o controlo negativo da técnica de imunohistoquímica para as regiões pós- e pré-sináptica, respetivamente. | 108 |
| Figura 16 – Imunofluorescência da proteína LARP4 linhas celulares 3T3 (A) e C2C12 (B). | 109 |

Figura 17 - Análise dos níveis de luciferase, na linha celular HEK293, após transfetar com os vetores de expressão GFP e GFP-LARP4 sob os genes repórter SMN-FL e Luciferase-SMN-3'UTR FL (A). (B) Gráficos que demonstram os níveis de expressão da LARP4 após interferência com siRNA LARP4. Análise dos níveis de luciferase (C) e dos níveis de mRNA de luciferase (D), na linha celular HEK293, após transfetar as células com Controlo Negativo siRNA e o siRNA LARP4 e avaliar o seu efeito nos genes repórter SMN-FL e Luciferase-SMN-3'UTR FL. Todos os dados estão expressos em média±desvio padrão de 3 amostras independentes. ns significa não significativa com $p > 0,05$ e * significa significância com $p < 0,05$, após aplicação do teste estatístico test t student.

.....111

Figura 18 - Análise funcional da interação entre LARP4 e Smn endógena, após transfetar as linhas celulares HEK293 e SK-N-Be(2)c com os vetores de expressão GFP e GFP-LARP4 e verificados os níveis de expressão através de PCR em tempo real (A) (B). Na linha celular HEK293 verificou-se que não há alteração nos níveis de expressão da proteína Smn endógena (B), nem dos níveis de expressão de mRNA do SMN endógeno (C). Na linha celular SK-N-Be(2)c observou-se um aumento de expressão da proteína Smn endógena (E), sem alteração nos níveis de expressão de mRNA do SMN endógeno (F). Todos os dados estão expressos em média±desvio padrão de 3 amostras independentes. ns significa não significativa e * significa significância com $p < 0,05$, após aplicação do teste estatístico test t student.

.....113

Figura 19 - Gráficos que demonstram os níveis de expressão da LARP4 após ensaios de interferência nas linhas celulares HEK293 e SK-N-Be(2)c114

Figura 20 - Análise funcional da interação entre LARP4 e Smn endógeno, após transfetar as linhas celulares HEK293 e SK-N-Be(2)c com siRNA controlo negativo e SiRNA LARP4, durante 72 horas. Na linha celular HEK293 verificou-se que não há alteração nos níveis de expressão da proteína Smn endógena (em cima, esquerda), mas uma diminuição dos níveis de expressão do mRNA do SMN endógeno (em baixo, esquerda). Na linha celular SK-N-Be(2)c não se observou alterações na expressão da proteína Smn endógena (em cima, direita), mas uma diminuição dos níveis de expressão de mRNA do SMN endógeno (em baixo, direita). Todos os dados estão expressos em média±desvio padrão de 3 amostras independentes. ns significa não significativa com $p > 0,05$ e * significa significância com $p < 0,05$, após aplicação do teste estatístico test t student.

.....115

Figura 21 - Diagrama do método de clonagem do construto pUAST_CG11505 (A). Imagens de eletroforese em geral da reacção de PCR do gene CG11505 (B) e da reacção de restrição do vetor pUAST e do produto de PCR do gene CG11505, com as enzimas de restrição NotI e XbaI (C).118

Figura 22 - Quantificação relativa da expressão da proteína CG11505 em moscas transgénicas. **A)** Western-Blot das amostras de proteína das moscas transgénicas balanceadas obtidas, identificando a expressão da proteína CG11505, que se encontra acoplado a um tag de HA, através do recurso do anticorpo anti-HA. **B)** Gráfico da quantificação da expressão da proteína CG11505 nas moscas transgénicas, tendo como controlo a proteína 3c11....120

Figura 23 - Resultados dos cruzamentos das diferentes moscas transgénicas com os drivers elav-GAL4 e GMR-Gal4 (A) e imagens de olhos das moscas transgénicas que têm impacto na retina dos olhos (B).....121

| | |
|---|-----|
| Figura 24 - <i>Imagens do fenótipo nos olhos das moscas controlo e da mosca CG11505-RNAi</i> | 122 |
|---|-----|

Índice de Tabelas

| | |
|--|----|
| Tabela 1 - Razões da vulnerabilidade dos neurónios motores a lesões. Adaptado de Ferraiuolo, Kirby, Grierson, Sendtner, & Shaw (2011) | 35 |
| Tabela 2 – <i>Características clínicas e moleculares dos vários fenótipos da SMA. Tabela adaptada de (Bowerman et al., 2017)</i> | 40 |
| Tabela 3 - <i>Identificação das proteínas selecionadas nos extratos proteicos HEK293 e SK-N-Be(2)c por espectrometria de massa e respetivos “scores” . Tabela adaptada de Oliveira (2010)</i> | 64 |
| Tabela 4 - <i>Vantagens e desvantagens do modelo animal Drosophila melanogaster. Tabela adaptada Prüßing et al. (2013)</i> | 68 |
| Tabela 5 - <i>Lista de primers utilizados nos ensaios de PCR em tempo real</i> | 90 |
| Tabela 6 - <i>Lista de primers utilizados no PCR convencional no processo de clonagem do vetor pUAST-CG11505</i> | 93 |
| Tabela 7 - <i>Lista de primers utilizados na sequenciação do vetor de expressão pUAST-CG11505</i> | 96 |

Lista de Abreviaturas

CO₂ – Dióxido de Carbono

DNA – Ácido Desoxirribonucleico

dsRNA – RNA de dupla cadeia

eIFs – Fatores de iniciação eucariota

FBS – *Fetal Bovine Serum*

GABA – Ácido Gama-aminobutírico

GAPDH – Gliceraldeído 3-Fosfato di-hidrogenase

LARP – *La-related protein*

LB – Líquido *Luria Broth*

miRNA – Micro RNA

mRNA – Ácido Ribonucleico mensageiro

mRNP – Ribonucleoproteína mensageira

NAIP – *Neuronal Apoptoses Inhibitory Protein* (Proteína Inibidora da Apoptose Neuronal)

ncRNA – Ácido Ribonucleico não codificante

PABP – *Poly(A) binding protein*

PBS – *Phosphate Buffer Saline*

PBS-T – PBS com *Tween 20*

PCR – Polimerase em cadeia de recção

REP – *Rough Eye Phenotype*

RISC - Complexo silenciador induzido por RNA

RBP – *RNA Binding Proteins* (Proteínas que se ligam ao RNA)

RRM – *RNA recognition motif* (Motivo de reconhecimento de RNA)

RNA – Ácido Ribonucleico

RNAi – RNA de interferência

SMA – *Spinal Muscular Atrophy* (Atrofia Muscular Espinal)

SMN – *Survival Motor Neuron*

siRNA – *Small interfering RNA*

snRNA – Pequeno RNA nuclear

snRNP – Partículas ribonucleoproteicas nucleares pequenas

TA – Temperatura ambiente

tRNA – RNA transferência

UAS – *Upstream Activating Sequences*

UTR – Região não traduzida

I. Introdução

1.1 Neurónio Motor: da neurobiologia à clínica

1.1.1 Biologia e Histologia do Neurónio Motor

Os tecidos e os órgãos do corpo são constituídos por células que, de acordo com a sua função específica e a sua interação, determinam a função do órgão. No sistema nervoso existem dois tipos principais de células: os neurónios e as células da glia, sendo os neurónios as únicas células responsáveis pela função do cérebro (Bear, Connors, & Paradiso, 2016; Kandel, Schwartz, Jessell, Siegelbaum, & Hudspeth, 2013). No cérebro adulto humano o número de neurónios é similar ao número de células da glia, cerca de 86 mil milhões de cada tipo (Azevedo et al., 2009). Os neurónios são as células do sistema nervoso responsáveis pela deteção das alterações do meio envolvente e por comunica-las entre si, de forma a direcionar as respostas às alterações detetadas. Por outro lado, as células da glia contribuem para a função cerebral através do isolamento, do apoio e da nutrição dos neurónios vizinhos (Bear et al., 2016).

O neurónio é constituído por quatro regiões morfológicamente bem definidas: o corpo celular ou soma, as dendrites, o axónio e os terminais pre-sinápticos, sendo que cada uma destas regiões tem um papel específico na geração dos sinais e na comunicação com os restantes neurónios (Kandel et al., 2013). Os neurónios são classificados de acordo com a sua estrutura, nomeadamente conforme a morfologia das neurites (dendrites e axónios) e das estruturas que inervam/função (Bear et al., 2016; Squire et al., 2008). Em função do número de neurites que emergem do soma, os neurónios podem ser classificados em unipolares, se apresentam apenas uma única neurite, e em bipolares ou multipolares, se apresentam duas ou mais neurites, respetivamente. Relativamente à sua função, os neurónios são classificados em neurónios sensoriais, neurónios motores e interneurónios (Squire et al., 2008).

A estrutura morfológica especializada do neurónio motor é bastante similar à dos restantes neurónios, apresentando um soma de onde emergem duas neurites constituídas por dendrites curtas e um único e longo axónio (Lambrechts, Robberecht, & Carmeliet, 2007). Os neurónios motores estão localizados em diversas regiões do sistema nervoso central e são classificados em dois tipos: os

neurónios motores superiores ou centrais e os neurónios motores inferiores ou periféricos (Stifani, 2014). No caso dos neurónios motores superiores, o seu corpo celular localiza-se na região pré-motora e motora primária do córtex cerebral, dando origem aos tratos corticoespinal e corticobulbar. Por outro lado, os neurónios motores inferiores apresentam o seu corpo celular nos núcleos dos nervos cranianos no tronco cerebral e na matéria cinzenta da medula espinal, mais especificamente ao nível do corno anterior (Stifani, 2014; Valetdinova, Medvedev, & Zakian, 2015).

Relativamente à relação inter-neuronal, os neurónios motores superiores recebem o seu *input* do córtex motor e essa informação é transmitida pelos seus axónios até aos neurónios motores inferiores. A sua comunicação é feita com recurso ao neurotransmissor de glutamato. A sinapse realizada pelos neurónios motores superiores acontece apenas no interior do sistema nervoso central. No caso dos neurónios motores inferiores, estes recebem informação dos neurónios motores superiores, dos neurónios sensoriais e dos interneurónios. O longo axónio dos neurónios motores, que é uma característica única e específica deste tipo de neurónio, realiza a sinapse no exterior do sistema nervoso central com recurso ao neurotransmissor acetilcolina, nomeadamente com as fibras musculares, dando origem à junção neuromuscular (Figura 1) (Stifani, 2014).

A junção neuromuscular é uma sinapse tripartida composta pelo axónio do neurónio motor (pré-sinapse), pelas fibras musculares (pós-sinapse) e pelas células da glia associadas à sinapse, que se designam por células de Schwann terminais. Esta sinapse é uma sinapse colinérgica que permite a transmissão de informação através de potenciais de ação entre os neurónios motores e as fibras musculares, com a função do controlo do movimento muscular (Burvill, 2009; Campanari, García-Ayllón, Ciura, Sáez-Valero, & Kabashi, 2016).

Os neurónios motores são células pós-mitóticas e o seu número diminuiu ao longo do tempo, devido a não apresentarem a capacidade de multiplicação. A suscetibilidade dos neurónios motores ao processo patológico resulta na génese das doenças do neurónio motor, não se conhecendo a razão da sua especificidade (Shaw, 1999).

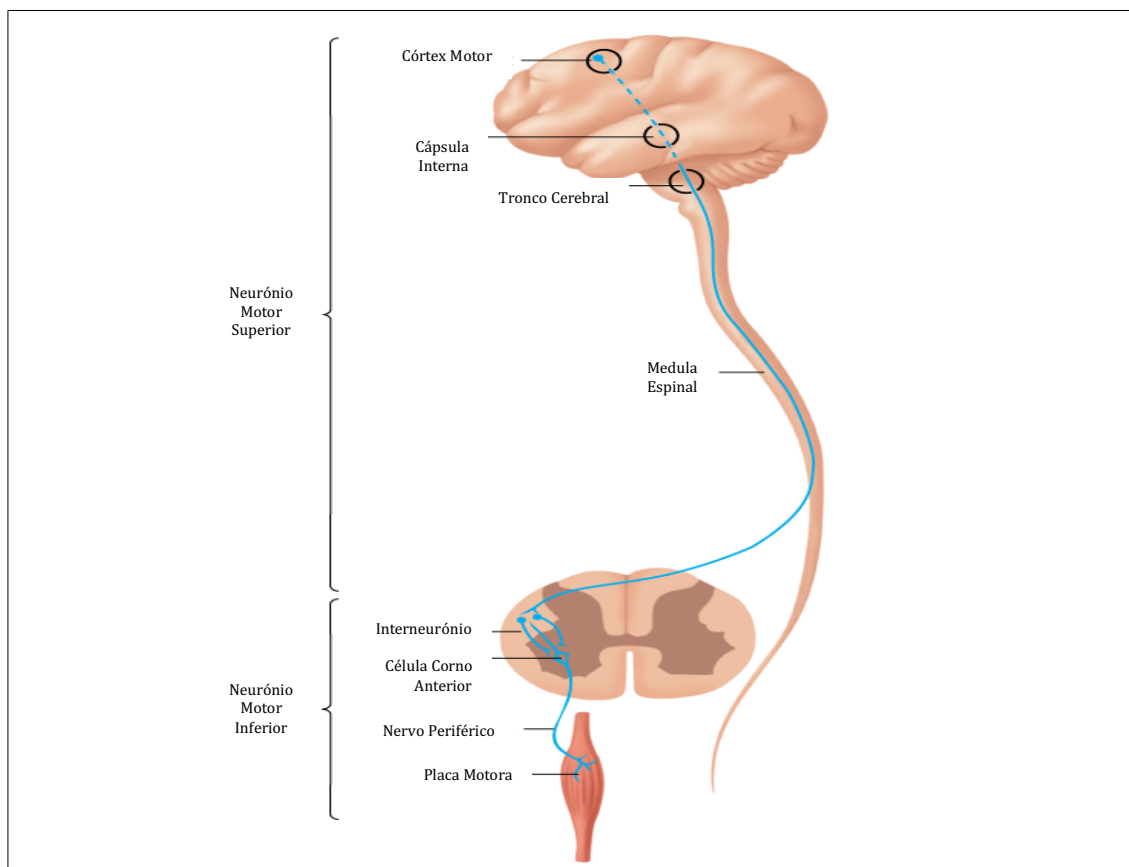


Figura 1 – Localização anatômica do neurônio motor superior e inferior e esquema do circuito motor. Imagem adaptada de (McKean, Ross, Dressler, & Scheurer, 2017)

Existe um conjunto de características específicas dos neurónios motores que os tornam vulneráveis. Uma dessas características é serem uma das células do sistema nervoso central com maiores dimensões, sendo responsáveis pela enervação das fibras musculares dos músculos dos membros inferiores a mais de um metro de distância. Para que isto seja possível, os neurónios motores necessitam não só de um elevado consumo de energia, requerendo dessa forma um maior número de mitocôndrias na sua constituição celular, mas também uma elevada estabilidade do citoesqueleto, possuindo na sua constituição proteínas neurofilamentares responsáveis por essa função. Outra característica particular dos neurónios motores é a presença, na sua superfície celular, de recetores de glutamato que quando ativados tornam os neurónios motores vulneráveis ao processo neurotóxico mediado pelo cálcio. Estes recetores de glutamato permeáveis ao cálcio são diferentes de outros presentes noutras células do sistema nervoso central. Este processo neurotóxico acontece porque os neurónios

motores não apresentam proteínas tamponadoras do cálcio intracelular (“*intracellular calcium buffering proteins*”) que protegem os neurónios do aumento excessivo deste ião (Shaw, 1999).

1.1.2 As Doenças do Neurónio Motor

1.1.2.1 *Epidemiologia das doenças do neurónio motor*

As doenças do neurónio motor são um grupo patológica e clinicamente heterogéneo que afeta primariamente os neurónios motores, responsáveis pelo controlo do movimento voluntário. Estas patologias são frequentemente responsáveis pela incapacidade e morte num período inferior a três anos após a manifestação dos primeiros sintomas.

A incidência destas doenças aumentou em cerca de 46% durante a década de 90, tendo atingido nessa época uma incidência de 1,9 por 100,000 pessoas/ano. Em contrapartida, nas décadas de 60-70, correspondia a uma incidência de 1,3 por 100,000 pessoas/ano (Worms, 2001). Atualmente, a incidência varia entre 0.6 e 2.4 por 100,000 pessoas/ano em diferentes população (Cronin, Hardiman, & Traynor, 2007). Considerando diferentes regiões do globo, observou-se um aumento de 162% nos países do sul da Europa, contrastando com o aumento de 17% no norte da Europa, não se observando alteração na América do Norte (Worms, 2001). Relativamente ao género e à idade dos doentes, verifica-se um maior número de casos no género feminino em relação ao género masculino e um pico de incidência entre os 75 e os 79 anos de idade, diminuindo posteriormente (Logroscino et al., 2008; Worms, 2001). No que toca à mortalidade, verificou-se um aumento na taxa de 1,2 por 100,000 pessoas/ano na década de 60-70 para 1,9 por 100,000 pessoas/ano na década de 90. Para estas alterações epidemiológicas têm contribuído o aumento da esperança média de vida, a definição dos critérios de diagnóstico *El escorial*, no início da década de 90, e a otimização da definição da classificação da patologia, segundo a classificação ICD, e na obtenção dos dados de óbitos (Worms, 2001).

1.1.2.2 Classificação e clínica das doenças do neurónio motor

As doenças do neurónio motor podem ser classificadas em hereditárias e esporádicas e ainda consoante o envolvimento do neurónio motor: inferior ou superior. As patologias esporádicas não são hereditárias e representam cerca de 90% dos casos das doenças do neurónio motor. A sua causa precisa não é conhecida, contudo pressupõem-se que esteja relacionada com fatores ambientais, como o exercício físico e a dieta, que juntamente com o fator genético tornam a pessoa suscetível à doença (Burvill, 2009; Filippi et al., 2010; Statland, Barohn, McVey, Katz, & Dimachkie, 2015).

O espectro clínico de classificação das doenças do neurónio motor pode variar desde uma entidade patológica que envolve apenas o neurónio motor superior (Atrofia Muscular Espinal, Atrofia Muscular EspinoBulbar e Atrofia Muscular Progressiva) ou o neurónio motor inferior (Esclerose Lateral Progressiva e Paraplegia Espástica) até ao envolvimento de ambos os tipos de neurónios (Esclerose Lateral Amiotrófica); esta última é responsável por cerca de 90% dos casos das doenças do neurónio motor (Filippi et al., 2010; Sendtner, 2005; Statland et al., 2015).

Clinicamente, os doentes apresentam uma heterogeneidade de sinais e sintomas, dependendo do neurónio motor afetado, sendo o tempo médio de diagnóstico de cerca de 14 meses. Porém, o fenótipo clínico e o local de início da doença tem implicações no seu desenvolvimento (Dharmadasa et al., 2017). De acordo com os mesmos autores, se houver uma lesão ao nível do neurónio motor superior, o doente apresenta um conjunto de sinais clínicos específicos, nomeadamente, espasticidade, movimentos incontroláveis e diminuição da sensibilidade à estimulação de reflexões superficiais. Por outro lado, caso seja o neurónio motor inferior afetado, há o desenvolvimento de paralisia da região enervado por esse neurónio (Figura 2).

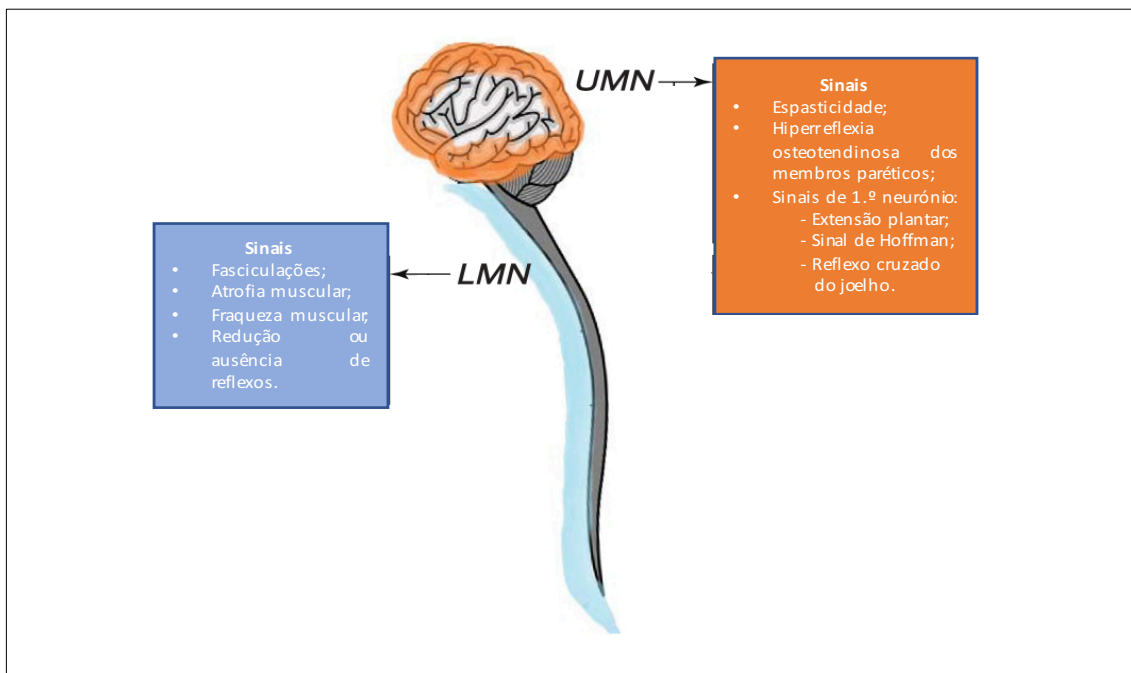


Figura 2 - Sinais associados a neurónios motores superiores (UMN) e inferiores (LMN) no desenvolvimento das Doenças do Neurónio Motor. Imagem adaptada de Dharmadasa et al. (2017)

1.1.2.3 Mecanismos fisiopatológicos das doenças do Neurónio Motor

Mann e Yates, nos anos 50, observaram uma série de alterações celulares em dois doentes: um de idade avançada, em que foi diagnosticada doença do neurónio motor, quatro meses após o início da sintomatologia, e faleceu devido a um acidente; e outro paciente com doença avançada que faleceu aos 52 anos de idade. Os autores analisaram um conjunto de secções da medula espinhal destes doentes. As observações realizadas concluíram que não havia diferença no rácio de RNA nuclear e nucleolar em relação ao citoplasma nas células de Purkinje entre os três indivíduos, bem como não havia diferença no tamanho do núcleo dos neurónios que se encontram no núcleo do trigémeo e do oculomotor. Mas nos neurónios de outros núcleos dos nervos cranianos observaram uma alteração do tamanho nuclear, que associaram a uma possível alteração degenerativa precoce. Para além destas observações, as observações mais notórias foram ao nível dos cornos anteriores da medula espinhal, mais acentuadas no estado avançado da doença.

Relativamente aos ácidos nucleicos, verificaram que no núcleo dos neurónios afetados havia a presença de aglomerados de heterocromatina e uma diminuição de RNA citoplasmático, principalmente no estado avançado. Com estas observações, Mann e Yates concluíram que quando os neurónios motores são afetados, ocorre a inativação do DNA nuclear e, por conseguinte, a diminuição da síntese do RNA mensageiro associado à diminuição do RNA do núcleo e, progressivamente, do citoplasma, conduzindo à morte celular (Press, 1975).

A marca patológica das doenças do neurónio motor é a morte seletiva dos neurónios motores, sendo a doença mais estudada até ao momento a Atrofia Lateral Amiotrófica (ALS). A razão pela qual este grupo específico de neurónios é suscetível a lesões, na presença de mutações que afetam proteínas ubíquas, não se encontra completamente compreendida. Porém, os neurónios motores apresentam um conjunto de características que podem predispor para a sua degeneração (Tabela 1).

| Premissas da vulnerabilidade dos neurónios motores à lesão | |
|---|---|
| ✓ | Células de grande tamanho, com axónios longos; |
| ✓ | Necessita de um mecanismo robusto de transporte axonal; |
| ✓ | Apresenta um consumo metabólico elevado, maior formação de espécies reativas de oxigénio e com tendência ao stress oxidativo |
| ✓ | Vulnerabilidade à excitotoxicidade e desregulação da homeostasia intracelular do cálcio: <ul style="list-style-type: none"> • Alta expressão de recetores AMPA permeáveis ao cálcio e sem a subunidade GluR2; • Baixa expressão de proteínas tamponantes de cálcio; • Dependente de mecanismos eficientes de transporte de recaptção do glutamato sináptico; |

Tabela 1 - Razões da vulnerabilidade dos neurónios motores a lesões. Adaptado de Ferraiuolo, Kirby, Grierson, Sendtner, & Shaw (2011)

Atualmente, a evidência científica sugere que a morte seletiva dos neurónios motores é multifatorial devido a uma complexa interação entre vários mecanismos, incluindo os fatores genéticos, stress oxidativo, excitotoxicidade, agregação de proteínas e danos em processos celular críticos (Lambrechts et al., 2007; Santa-Cruz & Tapia, 2014; P. J. Shaw, 1999; Shaw & Eggett, 2000). Porém, a

relevância destes fatores pode variar nos diferentes subgrupos de doentes que são afetados pelas doenças do neurónio motor (Shaw & Eggett, 2000).

As doenças do neurónio motor têm uma causa genética, havendo múltiplas alterações a nível de genes associados à degeneração dos neurónios motores (Shaw, 1999). Inicialmente, pensava-se que a cada entidade clínica estava associada uma mutação genética. Porém, com a evolução do conhecimento científico verificou-se que a mesma mutação genética pode ocorrer em entidades clínicas diferentes (Lambrechts et al., 2007). Como, por exemplo, a mutação do gene responsável pela produção da enzima Cu/Zn superóxido dismutase - *SOD1* – apresenta fenótipos distintos, em que podem se encontrar afetados de forma variável os neurónios motores e sensoriais, os tratos espinocerebelar e corticoespinal; sendo este gene responsável por 20% dos casos familiares de ALS (Lambrechts et al., 2007; Shaw, 1999). Para além disso, também se tem verificado que mutações em proteínas que se ligam ao RNA (*RNA binding proteins*), como a TDP43 ou o FUS, resultam numa alteração do metabolismo do RNA (transcrição, *splicing*, transporte, degradação ou tradução), apresentando um papel importante na patogénese da ALS (Bozzo, Mirra, & Carri, 2017).

Os efeitos de stress oxidativo em células pós-mitóticas, como o caso dos neurónios, podem ser cumulativos e a lesão celular devido a radicais livres é uma causa potencialmente importante relacionada com a idade na função neuronal que está associada às doenças neurodegenerativas. No caso da ALS, em que cerca de 20% dos casos são devidos à mutação do gene *SOD1*, que codifica uma proteína responsável pela defesa antioxidante celular, tem havido um particular interesse no estudo do stress oxidativo no desenvolvimento desta patologia (Shaw, 2005)

No sistema nervoso central, o principal neurotransmissor excitatório é o glutamato. Na fenda sináptica, este neurotransmissor é removido através de proteínas de transporte que se encontram localizadas nas células da glia perisinápticas (Shaw, 1999). O aumento da estimulação dos recetores de glutamato resulta na excitotoxicidade dos neurónios devido a alterações na homeostasia do cálcio intracelular e à produção excessiva de radicais livres. Esta alteração leva à morte dos neurónio motores que são particularmente suscetíveis à ativação dos recetores AMPA permeáveis ao cálcio (Santa-Cruz & Tapia, 2014; Shaw, 2005; Shaw, 1999). Para além disso, também o comprometimento da expressão e da

função da proteína transportadora de recaptação do Glutamato (EAT2), nas células da glia, resulta no aumento dos níveis do neurotransmissor na fenda sináptica. Nos casos de ALS em que há mutação do gene SOD1, ocorre uma maior sensibilidade dos neurónios motores à toxicidade pelo glutamato devido a alterações na expressão das subunidades dos recetores AMPA e à redução da expressão da EAAT2 (Shaw, 2005).

A mitocôndria é um organelo celular bastante importante quer na sobrevivência quer na morte dos neurónios motores devido à sua função nas vias metabólicas e de apoptose (Manfredi & Xu, 2005). Porém, alterações na função da mitocôndria predispõe à iniciação das vias intrínsecas de apoptose, ao aumento da geração de espécies reativas de oxigénio e ao aumento dos níveis de cálcio intracelular e por conseguinte na excitotoxicidade celular e morte dos neurónios motores (Manfredi & Xu, 2005; Santa-Cruz & Tapia, 2014).

A estrutura do citoesqueleto e a sua manutenção funcional revela uma grande importância na preservação do neurónio motor, visto apresentarem axónios compridos com comprimentos até um metro (Lambrechts et al., 2007; Shaw, 2005). O calibre dos axónios dos neurónios, a integridade celular e o transporte axonal estão dependentes da estrutura filamentosa constituída por neurofilamentos (Shaw, 2005). Mutações ou deleções nos seus diferentes subtipos resultam no desenvolvimento de agregados e na acumulação de proteínas no corpo celular dos neurónios motores e nos seus axónios proximais, levando ao desenvolvimento de diferentes doenças do neurónio motor como a ALS e a doença de Charcot-Marie Tooth (Lambrechts et al., 2007).

Os neurónios motores são altamente dependentes de um sistema de transporte intracelular eficiente com componentes anterógrado e retrógrado. O transporte axonal é realizado através da família das moléculas motoras cinesina (transporte axonal anterógrado) e dineína cinesina (transporte axonal retrógrado), que se movem ao longo dos microtúbulos. A mutação nos genes na família destas duas proteínas resulta na degeneração dos neurónios motores e no desenvolvimento de patologias do neurónio motor, como a Doença de Charcot-Marie-Tooth e a Paraplegia Espástica hereditária caso a mutação seja no gene *KIF1B*. Caso seja uma mutação no gene dinactina (*DCTN1*), que é responsável pela construção do

complexo de dineína e, por conseguinte, pelo transporte retrógrado resulta na doença neuronal motora progressiva (Lambrechts et al., 2007; P. Shaw, 2005). Portanto, o processo de degeneração do neurónio motor é complexo e multifatorial, sendo que o fator genético possa definir o desenvolvimento da degeneração. Porém ainda falta por esclarecer acerca dos fatores ambientes e genéticos que predis põem à forma esporádica das doenças do neurónio motor.

1.2 Atrofia Muscular Espinal

A Atrofia Muscular Espinal (*Spinal Muscular Atrophy* - SMA) é uma doença autossómica recessiva, sendo a doença do neurónio motor mais comum de caracter hereditário. Esta patologia é a causa de morte de caracter hereditário mais frequente em recém-nascidos e crianças (Iascone, Henderson, & Lee, 2015; Shababi, Lorson, & Rudnik-Schöneborn, 2014). Trata-se de uma doença rara, com uma incidência de cerca de 1:11.000 nados vivos, e com uma frequência de portadores de aproximadamente 1:54 (Sugarman et al., 2011).

1.2.1 Atrofia Muscular Espinal e sua clínica

A SMA foi descrita pela primeira vez por Werdnig, em 1891, após ter observado dois irmãos em idade pediátrica que apresentavam, desde os 10 meses de idade, fraqueza muscular na região proximal dos membros inferiores, de início progressivo. Entre 1893 e 1900, Hoffman descreveu um conjunto de sete casos em três famílias distintas, classificados como casos de severidade intermédia. Os primeiros dois casos de crianças que apresentavam paralisia flácida dos músculos do tórax e dos membros ao nascimento foram descritos por Sylvester e Beevor em 1899 e em 1903, respetivamente (Iascone et al., 2015).

1.2.1.1 Manifestações Clínicas

As manifestações clínicas da SMA resultam da perda progressiva dos neurónios motores localizados no corno anterior da medula espinal. Esta degenerescência dos neurónios motores resulta em fraqueza simétrica e atrofia dos músculos voluntários proximais dos membros inferiores e superiores e, possivelmente, de todo o tronco durante a evolução da doença. A degenerescência é mais acentuada nos neurónios motores responsáveis pela inervação dos músculos proximais relativamente aos músculos distais, para além de os membros inferiores serem mais fracos que os membros superiores.

Em 1991, no *International SMA Consortium Meeting*, a SMA foi formalmente classificada em três formas clínicas, aguda ou doença de Werdnig-Hoffmann (fenótipo tipo I ou severo), intermédia ou crónica infantil (fenótipo tipo II) e crónica juvenil ou doença de Kugelberg-Welander (fenótipo tipo III), de acordo com 3 critérios clínicos: idade de início da manifestação dos sintomas, o curso da doença e a esperança média de vida (Tabela 2) (Munsat & Davies, 1992).

O fenótipo de tipo I ou Doença de Werdnig-Hoffmann é o mais severo e o mais comum, correspondendo a 50% da população dos doentes com SMA. O início dos sintomas acontece antes dos seis meses de idade, com elevada mortalidade nos primeiros dois anos de vida. As crianças com o tipo severo apresentam um conjunto de manifestações clínicas como hipotonia, paralisa flácida simétrica, ausência do controlo do movimento da cabeça, não se conseguem sentar sem apoio e apresentam uma respiração paradoxal e o tronco em forma de sino devido ao enfraquecimento dos músculos intercostais sem alterações do diafragma. Para além destas manifestações características, podem também surgir fasciculações e fraqueza da língua devidas ao envolvimento bulbar. Estas alterações resultam no impedimento do desenvolvimento dos movimentos de chupar e de engolir, aumentando o risco de aspiração devido à diminuição da proteção aérea, sendo esta uma importante causa de mortalidade e morbilidade (D'Amico, Mercuri, Tiziano, & Bertini, 2011; Iascone et al., 2015; Lunn & Wang, 2008; Shababi et al., 2014).

| Fenótipos da SMA | | | | | |
|------------------|---|--|--|---|---|
| | 0 | I | II | III | IV |
| N.º cópias SMN2 | 1 | 2 | 3 | 3-5 | 3-5 |
| Idade de início | Útero | Maioria até aos 6 meses | 6-12 meses | Após os 18 meses (IIIa<3anos; IIIb>3anos) | Idade adulta |
| História Natural | - Morte peri-natal. | - 50% morre até aos 12 meses; - 90% morre até aos 24 meses sem ventilação invasiva. | - Esperança média de vida entre 30-50 anos dependendo da função respiratória. | - Esperança média de vida normal; - Perda de locomoção é variável; - Envolvimento respiratório incomum. | - Progressão lenta; - Locomoção mantida; - Esperança média de vida normal. |
| Clínica | - Perda motora generalizada; - Contraturas; - Incidência elevada de anomalias cardíacas congénitas. | - Hipotonia neonatal; - Fraca capacidade de alimentação e controlo da cabeça; - Insuficiência respiratória; - Não desenvolve capacidade de rebolar ou sentar sem ajuda. | - Capacidade para sentar sem apoio; - Incapacidade para andar; - Fraca musculatura respiratória. | - Caminhada sem assistência, mesmo que por breves instantes | - Diminuição da força proximal progressiva; - Predominância nos membros inferiores |

Tabela 2 – Características clínicas e moleculares dos vários fenótipos da SMA.

Tabela adaptada de (Bowerman et al., 2017)

Quando o início dos sintomas acontece entre os 7 e os 18 meses de idade, o fenótipo caracteriza-se como sendo de severidade intermédia ou tipo II. As crianças com este fenótipo conseguem sentar-se sem ajuda e algumas conseguem ficar de pé com recurso a suportes dos membros inferiores, mas não têm capacidade para andar de forma independente. A presença de tremores finos é característica deste fenótipo. Devido à fraqueza do reflexo de deglutição, poderá haver dificuldade em

manter ou aumentar o peso corporal. Como descrito no fenótipo anterior, a função respiratória é comprometida devido à fraqueza dos músculos intercostais e a tosse torna-se ineficaz, resultando na dificuldade da limpeza brônquica como consequência da desnervação bulbar. A insuficiência respiratória é a principal causa de morte durante a adolescência (D'Amico et al., 2011; Lunn & Wang, 2008). A doença de Kugelberg-Welander ou tipo III apresenta um quadro clínico heterogêneo, em que os sinais se manifestam após os 18 meses de idade, com uma esperança de vida normal. As crianças que apresentam este fenótipo conseguem desenvolver todas as marcas motoras do desenvolvimento, tendo a capacidade de andar de forma independente. Porém alguns doentes necessitam de cadeira de rodas durante a infância, enquanto que outros conseguem continuar a caminhar, bem como viver uma vida adulto produtiva, apresentando uma fraqueza muscular menor (Lunn & Wang, 2008; Russman, 2007).

Para além dos três fenótipos descritos anteriormente, foi identificado um novo fenótipo, o fenótipo tipo IV ou do adulto, em que a fraqueza geralmente surge na segunda ou terceira décadas de vida, com uma esperança de vida normal. Estes doentes apresentam um compromisso motor leve, sem nenhum problema respiratório ou nutricional e têm a capacidade de caminhar de forma independente, na idade adulta (D'Amico et al., 2011; Huang & Luo, 1983; Lunn & Wang, 2008; Russman, 2007).

1.2.1.2 Diagnóstico e meios complementares de diagnóstico

A SMA é uma doença de diagnóstico clínico difícil, visto apresentar uma baixa incidência e as manifestações clínicas, como hipotonia, parésia, arreflexia e fasciculações serem comuns a outras neuropatias. Contudo, por ser uma doença de progressão rápida, é fundamental que o diagnóstico seja realizado com a maior celeridade possível (Baioni & Ambiel, 2010).

O diagnóstico clínico da SMA é realizado com base na evidência de desnervação muscular detetada no estudo neurofisiológico e na biópsia muscular. O estudo neurofisiológico no diagnóstico clínico da SMA é realizado com base na evidência de desnervação muscular, sem qualquer impacto na condução dos nervos motores

e sensoriais. Para além disso, demonstra uma redução na amplitude dos potenciais de ação revelando uma redução da fibra motora e são observados potenciais de fibrilhação em repouso, nos casos de desnervação, podendo estar diminuídas a duração e a amplitude dos potenciais das unidades motoras e haver uma redução da velocidade de condução motora nos estadios mais precoces. Porém, importa realçar que este tipo de estudo nem sempre apresenta um padrão típico no início da doença (Baioni & Ambiel, 2010; Elsheikh & Kissel, 2014).

A biópsia muscular apresenta um conjunto de características histopatológicas características, como atrofia das fibras musculares (tipo I e tipo II), hipertrofia das fibras tipo I ou agrupamentos tipo fibra. Porém, estas características não são específicas da SMA, sendo observadas em outros casos de desnervação (Baioni & Ambiel, 2010). No caso da SMA tipo I podem observar-se grandes grupos de fibras atroficas intercalados com fascículos contendo fibras hipertrofiadas. A extensão da atrofia muscular é observada ao longo de todo o músculo (Elsheikh & Kissel, 2014).

Para além dos sinais clínicos anormais, a SMA pode ser acompanhada de alterações a nível analítico, nomeadamente nos valores da creatinina quinase, que podem apresentar-se cinco vezes baixo do valor normal (Baioni & Ambiel, 2010). Contudo, quando há suspeita da patologia, o teste genético é a primeira opção no diagnóstico laboratorial, não só por ser o teste que estabelece o diagnóstico definitivo, mas também porque evita recorrer-se a uma investigação com exames auxiliares de diagnóstico mais invasivos. A deleção homozigótica ou a conversão do exão 7 do gene *Survival of Motor Neuron* (SMN) 1 é detetada em 93% dos casos de SMA, com ou sem deleção do exão 8 (Baioni & Ambiel, 2010; Elsheikh & Kissel, 2014). Caso se suspeite que o doente tenha apenas uma cópia deste gene, devem ser investigadas a presença de outras mutações, inserções e deleções responsáveis pela disfunção homozigótica do gene (Baioni & Ambiel, 2010).

1.2.2 Bases moleculares da Atrofia Muscular Espinal

A variabilidade dos diferentes fenótipos da SMA resulta diretamente do background genético da doença. Em 1990, com recurso a análises de *linkage*,

vários autores concluíram que o gene responsável por esta doença estaria localizado ao nível do cromossoma 5, mais especificamente na região 5q11.2-q13.3. Em 1995, Lefebvre e seus colaboradores identificaram o gene SMN1 como ausente ou mutado em mais de 98% dos indivíduos no *cohort* de doentes com SMA, apontando para outro fator associado ao desenvolvimento da doença. Estudos subsequentes viriam a demonstrar que o gene SMN está presente em duas cópias quase idênticas no genoma humano, na mesma região do cromossoma 5, sendo esta característica única da espécie em comparação com os restantes genomas de mamíferos até hoje caracterizados, em que há apenas uma única cópia deste gene (Lefebvre et al., 1995; Tiziano, Melki, & Simard, 2013). O gene SMN é composto por 9 exões, denominados 1, 2a, 2b, 3, 4, 5, 6, 7 e 8, compreendendo uma região de cerca de 20kb (Burglen et al., 1996). Uma das cópias, o SMN1, localiza-se no lado telomérico do cromossoma 5. A segunda cópia, o SMN2, apresenta uma localização centromérica no mesmo cromossoma (Lefebvre et al., 1995; Tiziano et al., 2013).

Os genes SMN1 e o SMN2 são quase idênticos (99.9%) e codificam a mesma proteína, designada por *Survival of Motor Neuron* (SMN) (Lefebvre et al., 1995). A única diferença entre estes dois genes é a alteração silenciosa de 5 nucleótidos, que se encontram localizadas nas seguintes regiões: exão 7 (c.840C>T), exão 8 (nt 27869 G>A), intrão 6 (nt 27092 G>A) e intrão 7 (nt 27289 A>G e nt 27404 A>G) (Burglen et al., 1996; Lefebvre et al., 1995; Lorson, Hahnen, Androphy, & Wirth, 1999). Apesar de a alteração ser silenciosa, a transição dos nucleótidos na posição +6 do exão 7 (c.840C>T) resulta na alteração de uma sequência reguladora do *splicing*, um elemento *cis* que promove a inclusão do exão, levando à perda da interação com o regulador de *splicing* ASF/SF2 (Cartegni, Hastings, Calarco, de Stanchina, & Krainer, 2006). Esta substituição parece levar ainda a uma melhor ligação da proteína hnRNPA1, resultando assim num efeito negativo no *splicing* devido à criação de um silenciador de *splicing* de exões (Kashima & Manley, 2003; Lorson et al., 1999). Esta diferença funcional no gene SMN2 resulta numa diminuição da incorporação do exão 7 no transcrito final, em cerca de 90% dos transcritos. A tradução deste transcrito final resulta na produção de uma proteína truncada e incapaz de se oligomerizar, instável e que é rapidamente degradada (Figura 3) (Burghes & Beattie, 2009). Contudo, os restantes 10% dos transcritos

finais incorporam o exão 7 e, por conseguinte, resultam na tradução de uma proteína SMN completa e funcional.

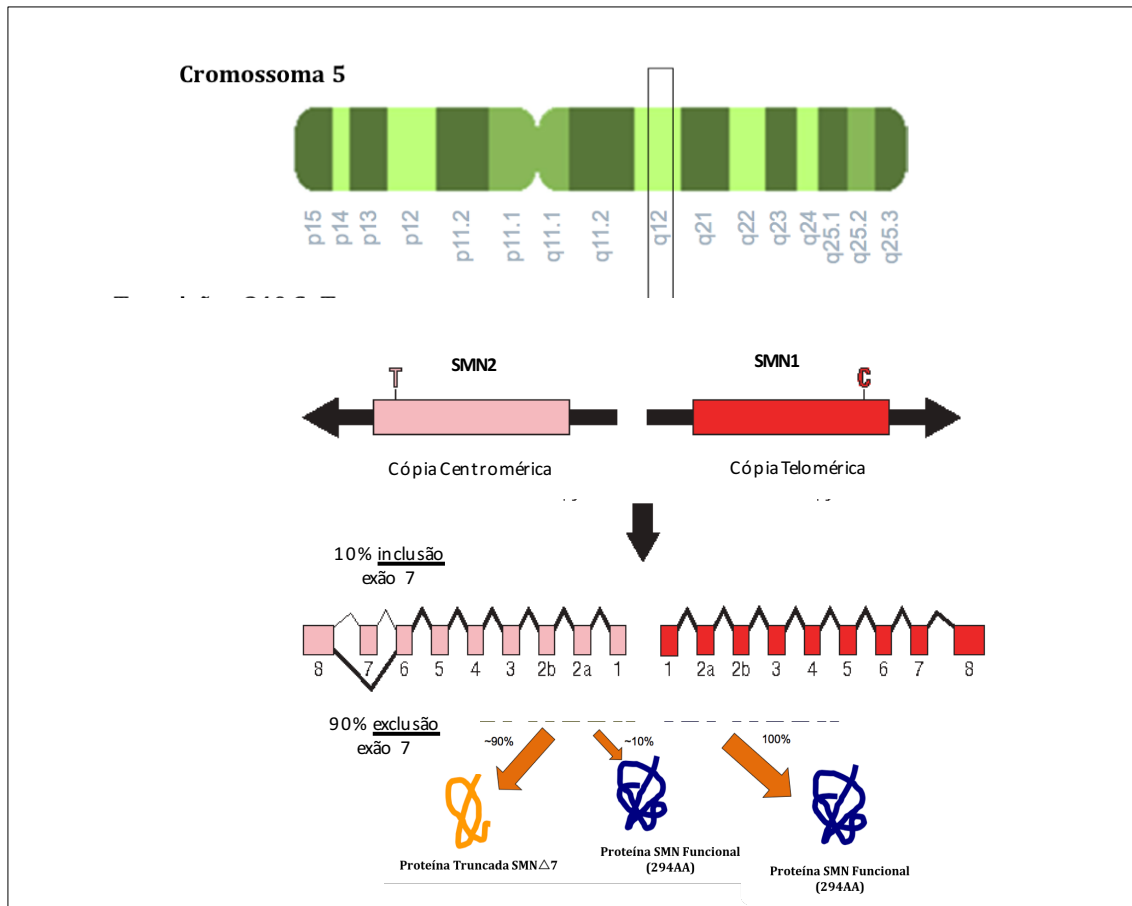


Figura 3 - Diagrama dos genes humanos SMN1 e SMN2 no cromossoma 5. Os doentes com SMA apresentam deleção ou mutação do gene SMN1 que é responsável pela produção da proteína SMN. O gene SMN2, devido a uma transição (c840 C>T) no exão 7, resulta na produção maioritária de uma proteína SMN truncada não funcional, rapidamente degradada. Esquema adaptado de Kolb & Kissel (2011) e Sumner, Paushkin, & Ko (2017)

1.2.2.1 Modificador da severidade da SMA: SMN2

A região codificante do SMN1 e do SMN2 são bastante idênticas, como referenciado anteriormente, apresentando apenas uma transição nucleotídica no

exão 7 (C>T) que é responsável pelo *splicing* alternativo e, como consequência, no desenvolvimento da patologia SMA (Lorson et al., 1999).

A transição nucleotídica no exão 7, referida anteriormente, encontra-se localizada no meio da sequência responsável pela regulação da inclusão do exão 7 nos transcritos do SMN designada por Intensificador de *Splicing* dos exões (ESS). Portanto, o gene do SMN1 apresenta o nucleótido Citosina (C) nessa posição promovendo a inclusão do exão 7 no transcrito final do SMN1 e, consequentemente, na produção da proteína SMN completa e funcional. Por outro lado, no gene SMN2 em que há o nucleótido Timina (T), nessa posição, perturba a sequência do ESS resultando na exclusão do exão 7 e, por conseguinte, na produção de um transcrito final sem o exão 7. Desta forma, há tradução de uma proteína parcial, não funcional que é incapaz de se oligomerizar e instável que é rapidamente degradada (Butchbach, 2016; Lorson & Androphy, 2000; Lorson et al., 1999). Para além disso Gabanella e seus colaboradores (2007) verificaram que o comprometimento na produção das snRNPs está relacionado com a severidade da doença, sugerindo que os transcritos necessários para a atividade dos neurónios motores são afetados por uma deficiência desta proteína.

Taylor e seus colaboradores (1998) referem que todos os doentes com SMA apresentam ausência homozigótica de SMN1 e nenhum apresenta deleção homozigótica do gene SMN2. Para além disso, os autores referem que a diminuição do número de cópias do gene SMN2 está associada ao aparecimento da sintomatologia mais cedo e um pior prognóstico. Assim, os doentes que apresentam quadros de maior severidade (fenótipo tipo I) possuem uma ou duas cópias de SMN2 em comparação com os de média severidade (fenótipo tipo II e III), que apresentam pelo menos três cópias do gene SMN2 (Figura 4). O número de cópias de SMN2 é assim um fator preditivo de severidade da doença, estando relacionado com aumento da sobrevivência dos indivíduos (Feldkötter, Schwarzer, Wirth, Wienker, & Wirth, 2002; Harada et al., 2002; Mailman et al., 2002; Watihayati et al., 2009).

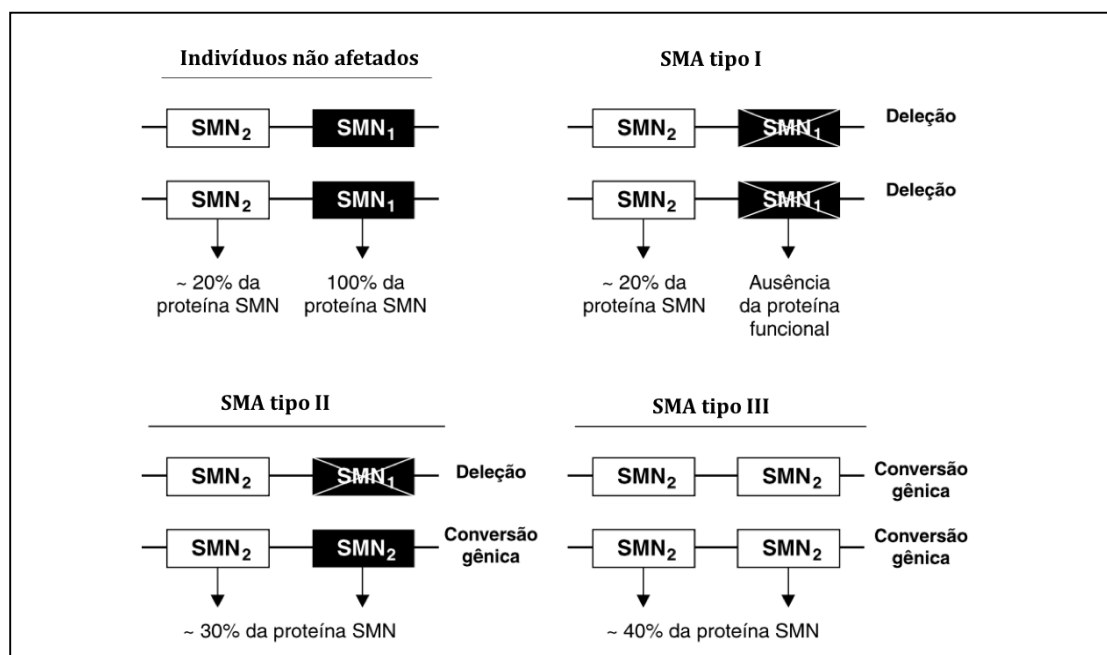


Figura 4 - Genótipo e correlação do número de cópias do gene SMN2 com os tipos de SMA. Imagem adaptada de Baioni & Ambiel (2010)

1.2.2.2 Proteína SMN e suas funções

Os genes SMN1 e SMN2 traduzem uma proteína ubíqua e idêntica, designada por SMN, que é constituída por 294 aminoácidos, tendo um peso molecular de 38 kDa (Lefebvre et al., 1995). Esta proteína encontra-se presente nos diferentes tecidos adultos com diferentes níveis de expressão. A sua maior expressão acontece no sistema nervoso central, nos rins e no fígado, apresentando uma menor expressão em fibroblastos e linfócitos. No tecido muscular esquelético e cardíaco a proteína SMN encontra-se expressa de forma moderada. Especificamente no sistema nervoso central, a proteína SMN encontra-se em elevada abundância na medula espinal, afetando de forma crítica os neurónios motores quando há uma diminuição significativa da quantidade desta proteína nos doentes com SMA tipo I (Battaglia, Princivalle, Forti, Lizier, & Zeviani, 1997; Coover et al., 1997). Porém não se sabe qual a razão de afetar apenas os neurónios motores quando há diminuição da sua expressão.

A nível celular, a proteína SMN encontra-se localizada em todas as células no núcleo e no citoplasma. No núcleo, a proteína encontra-se localizada em

estruturas sub-nucleares designadas por corpos de Cajal ou “*coiled bodies*” e nos *gems* ou “*gemini of coiled bodies*” (Coovert et al., 1997; Liu & Dreyfuss, 1996). Primariamente, foram identificados os *gems*, em células HeLa, como estruturas nucleares que contém uma elevada concentração de SMN (Liu & Dreyfuss, 1996). Os *gems* são similares no tamanho e em número aos corpos de Cajal, estando intimamente relacionados e apresentam um comportamento dinâmico com os corpos de Cajal devido, em parte, à capacidade de interação entre as proteínas Coilina e SMN (Hebert, Szymczyk, Shpargel, & Gregory Matera, 2001; Liu & Dreyfuss, 1996). Para além disso, os *gems* encontram-se fundidos com os corpos de Cajal no núcleo da maioria das linhas celulares e dos tecidos adultos. Contudo, as mesmas estruturas encontram-se separadas nos tecidos embrionários. Portanto, estas duas estruturas nucleares são distintas mas apresentam uma relação funcional dinâmica (Paushkin, Gubitz, Massenet, & Dreyfuss, 2002; Pellizzoni, 2007). Os corpos de Cajal contêm elevadas concentrações de fatores que estão envolvidos no processamento e no metabolismo do RNA, estando o seu número (por célula) inversamente correlacionado com a gravidade da doença (Gubitz, Feng, & Dreyfuss, 2004; Patrizi et al., 1999; Paushkin et al., 2002; Pellizzoni, Kataoka, Charroux, & Dreyfuss, 1998).

A proteína SMN, no núcleo, oligomeriza e forma um complexo estável, designado por complexo SMN. A oligomerização do SMN necessita, no terminal carboxilo, do domínio de oligomerização *YG box*, que se encontra interrompido nos casos de SMA devido a mutações no gene SMN2 (Li, Tisdale, Lotti, & Pellizzoni, 2014a).

O complexo SMN resulta da associação da proteína SMN com um conjunto de proteínas, conhecidas por Gemins: Gemin2 (SIP1), Gemin3/DP103 (DEAD-box RNA helicase), Gemin4, Gemin5/p175 (proteína repetida WD), Gemin6, Gemin7 e Gemin8 (Gubitz et al., 2004; Kolb, Battle, & Dreyfuss, 2007; Li, Tisdale, Lotti, & Pellizzoni, 2014). Apenas os Gemins 2, 3, 5 e 8 interagem diretamente com o SMN, os restantes interagem de forma indireta com o SMN (Gubitz et al., 2004; Li et al., 2014b). Estas proteínas apresentam uma co-localização com a proteína SMN nos *gems*, para além da sua presença no citoplasma; exceto a proteína Gemin4 que também se localiza no nucléolo (Charroux et al., 2000).

Como referido anteriormente, a proteína SMN encontra-se também como parte integrante do complexo SMN, no citoplasma, juntamente com as proteínas Gemin

2 a 8 e Unrip2 (Pellizzoni, 2007). Este complexo interage com várias proteínas de diversos complexos ribonucleoproteicos, funcionando como uma máquina de montagem dos mesmos (C. J. Sumner, 2007).

A principal função do complexo SMN é a montagem e a regulação da estabilidade dos complexos RNA-proteína das partículas ribonucleoproteicas nucleares pequenas (snRNPs), que são um componente essencial do spliceossoma. Este grande complexo é responsável pela catálise do *splicing* do pré-mRNA em todas as células (Pellizzoni et al., 1998; Will & Lührmann, 2001; Yong, Golembe, Battle, Pellizzoni, & Dreyfuss, 2004).

A proteína truncada produzida pelo gene SMN2 raramente é detetada em células ou tecidos de doentes ou de modelos animais com SMA, demonstrando que esta proteína é altamente instável, provavelmente devido à incapacidade de se oligomerizar e de se associar a outras proteínas por ausência do exão 7, que codifica o domínio *YG Box* pela oligomerização (Kostova et al., 2007; Li et al., 2014b; Lorson et al., 1999; C. J. Sumner, 2007).

1.2.2.3 Proteína SMN e o Sistema nervoso central

Nas células neuronais, a proteína SMN encontra-se localizada nas dendrites e nos axónios e ao nível do cone de crescimento dos neurónios motores. Ao nível do axónio, a proteína SMN concentra-se em grânulos e estes exibem movimentos bidirecionais rápidos dependentes dos microtúbulos e dos microfilamentos (Béchade et al., 1999; Zhang et al., 2003). Rossoll e seus colaboradores (2003) observaram a interação da proteína SMN com a hnRNP R, formando o complexo hnRNP R-Smn, que interage com a região 3'UTR do mRNA da β -actina, sendo necessário para o transporte eficiente do mRNA e no desenvolvimento dos cones de crescimento dos neurónios motores. Em modelos animais com deficiência da proteína SMN, nomeadamente em ratinho e peixe zebra, foi demonstrada a presença de axónios curtos e cones de crescimentos mais pequenos associados à deficiência de mRNA da β -actina, bem como um déficit específico no crescimento dos axónios nos neurónios motores (McWhorter, Monani, Burghes, & Beattie, 2003; Rossoll et al., 2003).

1.2.3 Modelos Animais

O conhecimento que atualmente se tem acerca da fisiopatologia da doença SMA é devido em grande parte a estudos realizados em modelos animais. Como foi referido anteriormente, o genoma humano contém o gene SMN1, sendo que a sua deleção ou mutação resulta na diminuição da proteína SMN e no aparecimento do fenótipo da patologia SMA. Este fenótipo pode ser diferente consoante o número de cópias do gene SMN2, presente apenas no genoma humano. Nos restantes organismos apenas existe um gene SMN, que é homólogo ao SMN1 dos seres humanos (Burghes & Beatiie, 2009; Rochette, Gilbert, & Simard, 2001). Por isso, a sua deleção leva à letalidade no período embrionário. Para que seja possível o estudo da SMA em modelos animais, foi necessário a introdução de uma cópia do gene SMN2 humano para que seja possível a reprodutibilidade do perfil genético e fenótipo da SMA. Em alternativa, foram construídos modelos animais com *knock-down* da expressão da proteína SMA por RNA de interferência, nalguns casos usando sistemas de expressão específicos de tecido muscular ou neuronal, que também simulam de forma eficaz o fenótipo da doença humana.

O genoma de *Caenorhabditis elegans* contém um único gene SMN homólogo ao gene SMN1 humano – *smn-1*, e que se traduz numa proteína idêntica em 36% à proteína SMN humana. A inativação deste resulta na morte das *C. elegans* no estadio larvar, o que implica que o gene *smn-1* é essencial à sobrevivência deste animal. O desenvolvimento de um animal em com uma deleção de uma região de 1Kbp que cobre a quase totalidade do gene *smn-1*, removendo todos os potenciais codões de iniciação (*smn-1(ok355)*, localizado no cromossoma I) resulta na diminuição da esperança média de vida, na perda progressiva das funções motoras, na diminuição da locomoção e atividade faríngea e num atraso no desenvolvimento. Este fenótipo é parcialmente compensado pela expressão de SMN nos neurónios, sendo que a sua expressão no tecido muscular não surtiu nenhum efeito (Briese et al., 2009). O modelo animal *C. elegans* com a mutação *smn-1(OK355)* é considerado um modelo eficiente para o desenvolvimento de *screens* em larga escala, no teste de drogas e na descoberta de modificadores do SMN, visto que são facilmente observadas alterações no fenótipo (Briese et al., 2009; Edens, Ajroud-Driss, Ma, & Ma, 2015).

Assim como em *C. elegans*, o genoma da *Drosophila* também contém apenas uma cópia do gene SMN, apresentando uma homologia de 41% da sua sequência relativamente ao genoma humano, sendo essencial à sua sobrevivência. Para além disso, em animais que apresentam mutações pontuais no gene *Smn*, semelhantes às que se encontram nos doentes de SMA, observa-se uma redução das correntes excitatórias pós-sinápticas, desorganização dos neurónios motores, nomeadamente uma arborização axonal defeituosa; a perda dos recetores de glutamato nas junções neuromusculares, a falha na formação dos filamentos finos dos músculos e o comprometimento das habilidades motoras. Estes resultados demonstram que o *Smn* tem um papel essencial na regulação do neurónio motor e nas funções neuromusculares na mosca *Drosophila* (Chan, 2003; Edens et al., 2015; Rajendra et al., 2007).

O modelo animal de *Drosophila* com mutação no gene *Smn* tem sido utilizado na identificação de genes que interagem com o SMN tendo identificado membros das vias de sinalização BMP (Chang et al., 2008) assim como genes que se encontram envolvidos na endocitose e no processamento do RNA (Hu et al., 2011). Sen e seus colegas (2011) verificaram que o *Smn* influencia a via de sinalização do Fator de Crescimento do Fibroblasto (FGF), quer através dos níveis de expressão dos seus componentes quer da própria via de sinalização. Para além disso, também observaram que a indução da via de sinalização do FGF pode restaurar os defeitos ao nível da junção neuromuscular que se encontram associados ao *Smn* nos modelos animais de *Drosophila melanogaster* em que há diminuição dos níveis de *Smn*. Desta forma, a via de sinalização do FGF pode atuar como um modificador de proteção da patologia SMA.

Sen e seus colaboradores (2013) tentaram definir o interatoma do *Smn* no modelo animal de *Drosophila melanogaster* e descobriram que muitos modificadores genéticos identificados possuem funções que foram previamente ligadas à atividade do *Smn*: metabolismo do RNA (*splicing* do RNA e ligação ao mRNA), controlo da tradução (complexo eIF3), endocitose (complexo Snap/SNARE) e transporte de proteínas. Contudo, alguns dos modificadores identificados não se conhece a sua associação à função do *Smn*.

O modelo animal de peixe zebra – *Danio rerio*, é considerado um excelente modelo na área da neurogenética devido ao facto de apresentar um sistema nervoso

simplificado, uma facilidade no desenvolvimento de animais transgênicos e na realização de testes genéticos. Para além disso, apresenta um conjunto de características que são importantes no estudo das doenças do neurónio motor e defeitos nas junções neuromusculares (Edens et al., 2015). McWhorter e seus colaboradores (2003) desenvolveram o método que induz a inibição do gene *Smn* através da injeção de *morpholinos antisense* em embriões em fase celular, por meio de ligação ao mRNA ou pré-mRNA, interferindo na tradução ou no splicing do *Smn*. Este tratamento com *morpholinos* resultou em defeitos no desenvolvimento dos neurónios motores, com fenótipos semelhantes aos dos doentes com SMA e que conduzem à desnervação. Para além do desenvolvimento deste método, também foram criados três mutantes: *smnY267stop* e *smnL265stop* que correspondem a mutações sem sentido ou *nonsense* que resultam na perda do exão 7 e, por conseguinte, a produção de uma proteína truncada e instável; e mutação missense do exão 7 (*smnG264D*) que corresponde à mutação *SMNG279V*, no genoma humano, que é a causa de SMA. Todas estas mutações resultam na redução e/ou na instabilidade da proteína *Smn* (Boon et al., 2009).

Tal como nos modelos animais anteriormente descritos, o genoma de ratinho também apresenta uma única cópia do gene *Survival Motor Neuron* – o *Smn1*. A sua inativação resulta na letalidade do animal devido à morte celular maciça que ocorre durante a formação inicial do blastocisto embrionário; sugerindo que a proteína SMN desempenha funções biológicas essenciais neste modelo animal (Edens et al., 2015; Schrank et al., 1997). Desta forma, é necessário a utilização de alelos condicionais para estudar a necessidade específica dos tecidos ao *Smn1*, permitindo a sobrevivência do animal. Portanto, Frugier e seus colaboradores (2000) realizaram a inativação do *Smn1* direcionado apenas ao tecido neuronal e demonstraram que ocorria perda dos neurónios motores assim com dos seus axónios. Para além disso, também realizaram a inativação do gene direcionada ao tecido muscular tendo-se observado uma contínua necrose muscular com fenótipo distrófico que levará à paralisia e à morte (Edens et al., 2015; Frugier et al., 2000). Portanto, estas observações sugerem que poderá ocorrer um envolvimento primário do musculo esquelético e, por conseguinte, contribuir diretamente para os defeitos motores observados, assim como para a degeneração do neurónio motor (Edens et al., 2015).

Para analisar os efeitos da diminuição da expressão da proteína SMN através da inativação do gene *Smn1*, foram desenvolvidos diferentes modelos animais com diferente número de cópias do gene SMN2 humano, resultando em modelos animais com os diferentes fenótipos de SMA (Park, Kariya, & Monani, 2010; Schrank et al., 1997). Atualmente existem um conjunto modelos animais que são utilizados no estudo da SMA: *Smn1*^{-/-};*SMN2*^{tg/tg};*SMNΔ7*^{tg/tg} (Le et al., 2005), *Smn1*^{hung}^{-/-};*SMN2*^{Hung}^{tg/tg} (Hsieh-Li et al., 2000), *Smn1*^{-/-};*SMN2*^{tg/tg};*SMN1A2G* (Monani et al., 2003) e *smn1*^{-/-};*SMN2*(566) (Monani, Coover, & Burghes, 2000)(Monani, Sendtner, et al., 2000). Os diferentes modelos animais referenciados anteriormente apresentam diferentes fenótipos, sendo que os dois primeiros são similares ao fenótipo II da patologia SMA humana, apresentando cerca de 13 dias de esperança média de vida. Por outro lado, os outros dois modelos animais apresentam um fenótipo similar ao fenótipo tipo III da SMA, apresentando o último modelo referido um fenótipo mais suave que o terceiro, em sintonia com o facto de apresentar um maior de cópias de SMN2.

1.3 Proteínas de ligação ao RNA (*RNA binding proteins*/RBP) e as doenças neurodegenerativas

O Dogma Central da Biologia Molecular foi descrito, em 1970, por Francis Crick, como um fluxo de informação entre três classes de biopolímeros: DNA, RNA e proteínas. O dogma descreve que a informação que se encontra numa sequência específica de DNA (gene) é transferida para uma molécula de RNA, servindo como molde para a síntese de um produto final funcional (proteína) (Crick, 1970). A transferência de informação contida no DNA para uma molécula de RNA ocorre por um processo designado por transcrição e a molécula sintetizada é designada por transcrito primário. Ao mesmo tempo que ocorre o processo de transcrição, o transcrito primário sofre modificações sucessivas resultando num transcrito secundário também designado por RNA maduro (Arraiano & Fialho, 2007; Bentley, 2014). Porém, nem todas as moléculas de RNA maduras seguem para a síntese proteica. Portanto, o genoma das células eucarióticas é transcrito em RNA mensageiro codificante (mRNA) e em diferentes espécies de RNA não codificante

(ncRNA), com funções na regulação da expressão génica, na função celular e na sobrevivência das células (Li et al., 2014b; Mignone, Gissi, Liuni, & Pesole, 2002; Nussbacher et al., 2015). A molécula de mRNA serve como molde para a síntese proteica, que se realiza nos ribossomas, através de um processo designado por tradução. Os dois processos encontram-se espacial e temporalmente separados nos espaços celulares, em que o processo de transcrição ocorre no núcleo da célula, ao contrário da tradução que se realiza no citoplasma (Arraiano & Fialho, 2007).

1.3.1 Regulação da expressão génica no processo de transcrição e de processamento do RNA mensageiro

A regulação da expressão génica inicia-se no núcleo da célula durante a transcrição de um gene. O gene começa a ser transcrito num local específico designado por local de início da transcrição que se encontra imediatamente precedido pela região promotora, onde a RNA polimerase se irá ligar.

A enzima RNA polimerase é responsável pela formação das ligações fosfodiéster entre os nucleótidos vizinhos, permitindo a formação da molécula de RNA. Contudo, nos eucariotas a transcrição envolve três enzimas RNA polimerases com funções diferentes, sendo elas a RNA polimerase I, II e III. A enzima que é responsável pela transcrição de RNA codificantes é a RNA polimerase II. Para além da transcrição do mRNA, esta polimerase também é responsável pela síntese de pequenos RNAs nucleares, com função no processamento do mRNA como os microRNAs (miRNA) e os ncRNA. A RNA polimerase I tem como função a transcrição de três RNAs ribossomais (28S, 18S e 5,8S) e a RNA polimerase III transcreve os genes que codificam os RNAs de transferência e a subunidade 5S do RNA ribossomal (Arraiano & Fialho, 2007; Eidem, Kugel, & Goodrich, 2016).

A transcrição de um gene segue um conjunto de 3 etapas: iniciação, elongação e terminação (Griffiths, Wessler, Lewontin, & Carroll, 2008). Para a RNA polimerase se ligar e iniciar o processo de transcrição são necessárias outras proteínas facilitadoras designadas por fatores gerais de transcrição. Portanto, a iniciação da transcrição ocorre após a construção de um complexo de pré-iniciação, que é

constituído pela RNA polimerase e pelos fatores gerais de transcrição (TFIIB, TFIID, TFIIE, TFIIF e TFIIH) na região promotora do DNA. A ligação do complexo proteico é definida pela sequência conservadora TATA, em alguns genes, que se encontra na região promotora e que irá sinalizar a região onde se irá ligar a polimerase (Arraiano & Fialho, 2007; Griffiths et al., 2008). Desta forma, os fatores gerais de transcrição cooperam com a RNA polimerase para se ligarem e abrirem a sequência promotora e iniciarem a síntese da molécula de RNA (Sainsbury, Bernecky, & Cramer, 2015).

Porém, a ligação apenas dos fatores gerais de transcrição produzem baixos níveis de atividade transcricional. Esta atividade pode ser aumentada ou diminuída com recurso a fatores específicos que se ligam a regiões do DNA designadas por potenciadores ou silenciadores, respetivamente, existindo cerca de 1400 no genoma humano. Um gene pode ter várias regiões potenciadores/silenciadores, podendo estas existir dentro ou fora da região do gene, algumas a milhares de nucleótidos longe da região do gene. A maioria dos fatores gerais de transcrição específicos para essas regiões e os fatores montados no promotor interagem através de um complexo mediador geral e um número de proteínas que não se ligam ao próprio DNA, designadas por cofatores. Embora os fatores gerais de transcrição e o complexo mediador sejam comuns à maquinaria transcricional de cada gene, os fatores gerais de transcrição específicos e os cofatores podem variar para cada gene (Kolovos, Knoch, Grosveld, Cook, & Papantonis, 2012). Desta forma, as flutuações na concentração de fatores gerais de transcrição específicos e os cofatores influenciam o tempo e a taxa de transcrição dos genes, proporcionando um mecanismo na regulação da expressão génica.

Após a RNA polimerase ser recrutada para a região do promotor do gene e formar o complexo de pré-iniciação com os fatores gerais de transcrição, ocorre o início da segunda etapa do processo de transcrição. Porém, muitos dos fatores gerais de transcrição são dissociados, mas parte do complexo de iniciação (TFIID, TFIIA e TFIIH) mantém-se associados com o promotor permitindo o início da transcrição de uma nova molécula de RNA (Kwak & Lis, 2013).

Quando ocorre a ligação da RNA polimerase ao promotor, a dupla hélice desenrola-se à frente da região do promotor, permitindo a formação de uma

cadeia de RNA complementar (Alberts et al., 2014). Esta etapa do processo de transcrição está dividida em duas partes: a elongação precoce e a elongação produtiva. Após o início da elongação, a enzima RNA polimerase II inicia a incorporação de nucleótidos na extremidade 3' com o intuito de estender o RNA nascente (Zhou, Li, Price, & David, 2012). A incorporação dos primeiros 10 nucleótidos no RNA nascente, a RNA polimerase II sai da etapa de iniciação do processo de transcrição. Porém, se a cadeia de RNA for menor que 10 nucleótidos, a RNA polimerase II ainda é considerada dentro da etapa de iniciação e encontra-se associada ao fator TFIIB e pode terminar prematuramente a nova molécula. Por outro lado, se tiver mais que 12 nucleótidos, o fator TFIIB é dissociado e o complexo de elongação é acoplado à RNA polimerase II de forma estável e inicia-se a elongação precoce (Kwak & Lis, 2013). Nos eucariotas superiores, durante a fase inicial da elongação, a RNA polimerase II transcreve os primeiros 20 a 60 nucleótidos iniciais na extremidade 3' da nova molécula de RNA e, posteriormente, realiza uma pausa entre o promotor e o primeiro nucleossoma (Kwak & Lis, 2013; Zhou et al., 2012). Desta forma, a elongação precoce pode ser definida como a transição da RNA polimerase II entre o promotor e a elongação produtiva, frequentemente acompanhada por um passo intermédio de pausa proximal do promotor. Durante esta pausa, são associados à polimerase os fatores de elongação primários NELF e DSIF, que permitem a estabilidade da enzima (Kwak & Lis, 2013; Zhou et al., 2012).

Após a transição para a segunda fase da elongação, a RNA polimerase II progride através do corpo do gene. Porém, para que seja possível a transição da polimerase para o modo produtivo, é necessário ocorrerem duas importantes modificações. Uma das modificações ocorre na extremidade 5' da nova molécula de RNA um processamento, designado por *capping*, que consiste na reação entre o 5' trifosfato do nucleótido terminal e o trifosfato de uma molécula de FTP, resultando numa ligação 5'-5'. Posteriormente, ocorre a adição de um grupo metilo à guanosina terminal, ocorrendo a transformação em 7-metilguanosina e a formação do *cap*. Este processamento protege o RNA da degradação, para além de ser essencial na exportação do mRNA para o citoplasma, através da regulação do complexo que se liga ao *Cap* (Arraiano & Fialho, 2007; Kwak & Lis, 2013; Ramanathan, Robb, & Chan, 2016; Zhou et al., 2012).

A outra modificação é a fosforilação do complexo de alongamento em que a enzima RNA polimerase II se encontra em pausa. A fosforilação do complexo é mediada pelo complexo de proteínas cinases P-TEFb, que também fosforilam os fatores de pausa NELF e DSIF permitindo a saída apenas do fator NELF do complexo de alongação. Nesse sentido, a RNA polimerase II e a molécula de RNA associada são modificadas para permitir o recrutamento de fatores adicionais de alongação e de processamento do RNA, com o intuito de ajudar a enzima polimerase a superar novas barreiras no corpo do gene assim como na transcrição e no processamento do RNA (Kwak & Lis, 2013; Zhou et al., 2012).

Finalmente, a terminação do processo de transcrição é um evento dependente do local poli(A), o que significa que o reconhecimento deste sítio precede ou acompanha a terminação e, portanto, é co-transcricional. Na maioria dos genes, mas não em todos, a clivagem no local poli(A) provavelmente fornece o ponto de entrada para a exonuclease XRN2, que age como um torpedo para promover a remoção da enzima polimerase do molde de DNA. Nesse sentido, tanto o reconhecimento quanto a clivagem do RNA no local poli(A) devem preceder a terminação (Bentley, 2014). Dessa forma, a RNA polimerase II transcreve a extremidade 3' do gene, onde a clivagem do RNA nascente e a poliadenilação definem o final do transcrito de mRNA. A terminação real de uma molécula transcrita pela RNA polimerase II ocorre em média 8Kb a jusante da extremidade 3' do gene em células humanas (Kwak & Lis, 2013).

A transcrição da RNA polimerase continua após o final do gene e sequências a jusante são removidas por um evento de processamento preciso que envolve a montagem de um complexo proteico grande guiado pelo sinal de consenso AAUAAA e sequências ricas em GU a montante do local de processamento. Este complexo multiproteico contém uma endonuclease que cliva o transcrito primário ou pre-mRNA a cerca de 10 a 35 nucleótidos a jusante do sinal de consenso para criar uma extremidade 3'. Para além disso, também contém uma polimerase poli(A) que adiciona um homopolímero de 200 a 300 resíduos de adenosina, formando uma cauda poli(A). Este último passo é coordenado entre um fator de splicing ligado ao último intrão do pre-mRNA e da polimerase poli(A), fazendo parte de uma das etapas do processamento do transcrito primário designa-se por poliadenilação e está relacionado com a estabilidade do mRNA bem como permite

a identificação de RNA codificantes e é um elemento essencial para a exportação do transcrito secundário para o citoplasma (Cooper & Hausman, 2013; Zhou et al., 2012).

A molécula de pré-mRNA é constituída por dois tipos de secções: os intrões e os exões. Os intrões são porções de gene que não codificam proteínas e são excluídas da sequência do mRNA funcional; enquanto que os exões são segmentos do gene que codificam proteínas que constituem a molécula madura de RNA (Alberts et al., 2014). Portanto, para a nova molécula de mRNA se tornar funcional é necessário a remoção dos intrões e a ligação dos exões vizinhos, através de um processo designado por *splicing*, que é exclusivo dos organismos eucariotas resultando numa molécula constituída por duas regiões não codificantes (5' UTR e 3'UTR) e uma região codificante (Figura 5) (Mignone et al., 2002). O processo de *splicing* é realizado com recurso ao spliceossomas que é constituído por complexos formadas por cinco pequenas ribonucleoproteínas nucleares (U1, U2, U4, U5 e U6), cada uma das quais é composta de uma molécula de pequeno RNA nuclear (snRNA), um conjunto de sete proteínas comuns e várias proteínas que são específicas para os snRNA (SJ Kolb, Sutton, & Schoenberg, 2010). O processo de *splicing* é facilitado pela interação com proteínas que se ligam ao RNA (*RNA binding protein* – RBP) (Fredericks, Cygan, Brown, Fairbrother, & Biology, 2015). Por vezes, o mesmo transcrito primário pode resultar em diferentes moléculas de mRNA através do processo de *splicing* alternativo. Este processo permite combinações diferentes dos exões que constituem o transcrito primário, resultando num mRNA diferente e, por conseguinte, na síntese proteica de diferentes proteínas. Desta forma, o *splicing* alternativo tem um papel importante na evolução, no desenvolvimento e na fisiopatologia de algumas doenças (Fredericks et al., 2015; SJ Kolb et al., 2010).

O processo de transcrição e de tradução ocorrem em espaços diferentes na célula, sendo necessário a exportação do mRNA do núcleo para o citoplasma para que ocorra a síntese proteica. Portanto, o transporte do núcleo para o citoplasma é essencial para a expressão génica nos eucariotas (Katahira, 2015).

O transcrito primário sofre um processo de maturação, de forma co-transcriptional, resultando numa molécula de mRNA funcional antes de ser transportado para o citoplasma através dos complexos de poros nucleares. No

decorrer desse processo, o transcrito associa-se a proteínas formando um ribonucleoproteico mensageiro (mRNP). Para além disso, proteínas adaptadoras do complexo TREX que se encontram envolvidas na exportação nuclear também se associam. Desta forma, a transcrição e o processamento do RNA e a exportação nuclear é um mecanismo importante visto que permite a maturação do transcrito primário assim como torna o mRNA maduro competente para o processo de tradução que ocorre no citoplasma. Portanto, estes processos nucleares permitem a precisão e a rapidez da expressão génica (Katahira, 2015).

1.3.2 Regulação da expressão génica no processo de tradução

A tradução é o processo em que a informação que se encontra na molécula de mRNA é traduzida numa proteína que representa uma molécula final e funcional. Portanto, neste processo, a molécula de mRNA é usada como molde, pelo ribossoma, para a síntese proteica. O processo de tradução, assim como a transcrição, é constituído por três etapas: iniciação, elongação e terminação. Para que seja possível a tradução da informação na molécula de mRNA em proteína, é necessário o ribossoma, RNA transferência (tRNA) e de pelo menos doze proteínas que são designadas por fatores de iniciação eucariota (eIFs) (Alberts et al., 2014; Cooper & Hausman, 2013).

Quando mRNA chega ao citoplasma, este é reconhecido pelo grupo de fatores eIF4 (A, B e G) que colabora no seu deslocamento até à subunidade 40S do ribossoma. Para além disso, a região *cap* do mRNA é reconhecido pelo eIF4E, que forma um complexo com os fatores eIF4A e eIF4G e este último fator também se liga à proteína PABP que se encontra associada á cauda poli(A) na região 3'UTR do mRNA. O tRNA, que contém o aminoácido metionina, liga-se à região P da subunidade 40S do ribossoma e este associa-se ao *cap*, formando o complexo de iniciação (Alberts et al., 2014; Cooper & Hausman, 2013).

Desta forma, o complexo de iniciação começa a percorrer a sequência do mRNA, no sentido 5'→3', até encontrar o codão de iniciação AUG e a ligação do tRNA de metionina a essa sequência. Após essa ligação, os fatores de iniciação dissociam-se do complexo e a subunidade 60S do ribossoma liga-se à pequena subunidade,

com recurso à eIF5B, formando um complexo ribossomal de iniciação 80S que permite o início do processo de tradução (Alberts et al., 2014; Cooper & Hausman, 2013).

O termino da síntese ocorre quando há o reconhecimento de um dos codões STOP (UAA, UAG ou UGA). O seu reconhecimento leva à ativação de fatores de libertação que terminam a síntese proteica após o estímulo da hidrólise da ligação entre o tRNA prévio e a cadeia polipeptídica formada e dissociação das subunidades do ribossoma permitindo a formação de um novo complexo de iniciação (Alberts et al., 2014).

Como foi explanado anteriormente, o processo de transcrição é o primeiro passo da expressão génica que pode ser controlado. Contudo, a modulação da expressão génica também pode ser controlada pela regulação da tradução da molécula de mRNA. A tradução de uma molécula de mRNA específica pode ser regulada por proteínas repressoras da tradução e por microRNA, que têm sido reconhecidos como reguladores centrais da expressão génica nas células eucariotas (Cooper & Hausman, 2013).

Um dos mecanismos de regulação da tradução é a ligação de proteínas repressoras - que são RBP - na região 5'UTR de mRNA específicos que bloqueiam o processo de tradução. Para além da região 5'UTR, a região 3'UTR também regula a tradução através da ligação de proteínas - que são RBP - a essa região do mRNA (Figura 5) (Alberts et al., 2014; Cooper & Hausman, 2013; Mignone et al., 2002).

Os repressores da tradução, maioritariamente, ligam-se ao fator de iniciação eIF4E e interferindo com a interação com o fator eIF4G, inibindo a tradução da síntese proteica. Para além disso, as proteínas que se ligam à região 3'UTR do mRNA também são responsáveis pela localização do mRNA em regiões específicas da célula, permitindo a produção da proteína em locais específicos da célula, sendo uma parte importante na regulação da tradução em vários tipos celulares (Cooper & Hausman, 2013; Mignone et al., 2002).

A modelação da atividade dos fatores de iniciação, particularmente do eIF2 e do eIF4E, também tem um efeito na tradução da molécula de mRNA. Principalmente, a regulação da atividade do fator eIF4E, que se liga ao *cap*, é outro ponto crítico em que os fatores de crescimento atuam no controlo da síntese proteica (Cooper & Hausman, 2013).

No decorrer das últimas décadas tem havido esforços no sentido de perceber o papel dos ncRNA na regulação da expressão génica, principalmente após a descoberta do RNA de interferência, em 1998. O processo de interferência do RNA é mediado por uma cadeia dupla de RNA que bloqueia a expressão génica ao nível da tradução. Desta forma, apresenta um papel importante na regulação pós-transcricional da expressão génica e na defesa inata contra a invasão de vírus. Os dois tipos mais comuns de pequenos RNAs, que medeiam estes processos, são os *small interfering RNA* (siRNAs) e os microRNA (miRNA); apresentando um comprimento de cerca de 22 nucleótidos e apresentam efeitos de silenciamento génico distintos (Cooper & Hausman, 2013; Lam, Chow, Zhang, & Leung, 2015). Em geral, o RNA de cadeia dupla (dsRNA) longa resulta de uma transcrição de genes celulares ou de patógenos infetantes, ou podem ser introduzidos na célula experimentalmente. O dsRNA é processado por uma enzima especialidade semelhante à ribonuclease III, designada por Dicer, no citoplasma em uma molécula de dsRNA menor conhecida por siRNA com 21 a 23 nucleótidos de comprimento e com 3 nucleótidos sobrepostos. O produto resultante interage e ativo o complexo silenciador induzido por RNA (RISC), resultando na clivagem da cadeia que se encontra no mesmo sentido do siRNA pela endonuclease argonaute 2 que é um dos componentes do RISC. A cadeia *anti-sense* irá permanecer ligada ao complexo RISC, ativando-o. Desta forma, o RISC ativo liga-se ao mRNA específico que é totalmente complementar à cadeia que está associada ao complexo, resultando no seu silenciamento (Cooper & Hausman, 2013; Lam et al., 2015).

No caso do miRNA, a sua transcrição é realizada pela RNA polimerase II, no núcleo, dando origem ao miRNA primário ou pri-miRNA que contém o *cap* a 5', a poliadenilação a 3' com uma estrutura *stem-loop* de cadeia dupla. Esta estrutura é clivada pela Drosha para formar o miRNA precursor ou pre-miRNA, que é um RNA de cadeia dupla que é constituído por 70 a 100 nucleótido com desemparelhamento intercalado e adota uma estrutura em forma de *loop*. O pre-miRNA é transportado para o citoplasma, pela Exportina 5, onde posteriormente é clivado pela Dicer num duplex de miRNA de 18 a 25 nucleótidos. De seguida, este associa-se aos RISC formando o complexo miRISC. Desta forma o duplex é desenrolado e a cadeia do mesmo sentido é descartada e não clivada como no caso

do siRNA. O miRNA de cadeia simples maduro guia o miRISC para os mRNAs alvo, ligando-se ao mRNA específica através de emparelhamento de bases complementares parciais, com a consequência de que o silenciamento do gene alvo ocorre via repressão da tradução, degradação e/ou clivagem do mRNA (Cooper & Hausman, 2013; Lam et al., 2015).

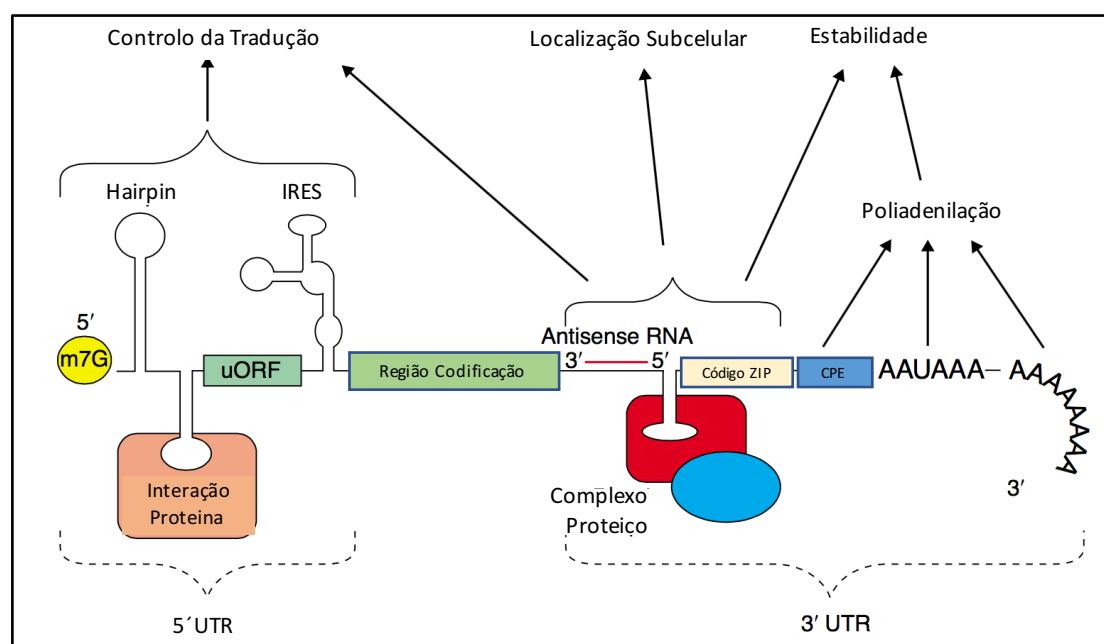


Figura 5 – Estrutura do mRNA de eucarióticas com exemplo de alguns elementos que afetam a expressão gênica. Imagem adaptada de Mignone et al. (2002)

1.3.3 Proteínas de ligação ao RNA (RNA Binding Proteins)

Tal como referenciado anteriormente, a biogénese e expressão dos mRNAs, também, depende em grande parte das interações com a família de RBP, formando complexos ribonucleoproteicos. Todas as RBPs ligam-se a moléculas de RNA (Glisovic, Bachorik, Yong, & Dreyfuss, 2008). A formação destes complexos envolve múltiplos passos, coordenados no espaço e no tempo. A evolução da complexidade dos organismos tem acompanhado a expansão do número de ncRNAs e RBPs e as mutações associadas a estas proteínas são muito frequentes em patologias humanas. Em particular, destacam-se as patologias do sistema nervoso central, causadas por mutações em genes expressos de forma ubíqua e que têm um papel no

processamento do RNA e na regulação pós-transcricional, como o SMN (Li et al., 2014b).

De acordo com Turner & Díaz-Muñoz (2018), existem cerca de 20.000 genes codificantes de proteínas nos mamíferos, sendo mais de 1.500 categorizado como RBPs e 692 como prováveis proteínas que se ligam ao mRNA. Normalmente, estas proteínas são modulares com domínios que medeiam funções adicionais. Estas proteínas ligam-se a sequências e/ou motivos estruturais do RNA através da combinação modular de um conjunto limitado de domínios de ligação ao RNA estruturalmente bem definido tais como o motivo de reconhecimento do RNA (RRM), domínio de homologia hnRNP K (KH), cadeia dupla RBP, domínio *Zinc-Finger* entre outros (Hentze, Castello, Schwarzl, & Preiss, 2018). O domínio RRM é o mais comum e muitos dos domínios RRM identificados interagem com proteínas e não com RNA. Por outro lado, o domínio KH liga-se quer à cadeia simples do DNA quer à do RNA e é ubíqua nos eucariotas, eubactéria e *archaea*. O RBP de cadeia dupla encontra-se quer nas bactérias quer nos eucariotas, interagindo com a cadeia dupla de RNA sem realizar contatos específicos nas nucleobases. Porém, o único domínio que se liga ao DNA nos eucariotas é domínio *Zinc-Finger*, mas também se liga ao RNA e foi primeiramente identificado no fator TFIIA (Hentze et al., 2018; Turner & Díaz-Muñoz, 2018).

As RBPs apresentam um conjunto de funções nos diversos aspetos da biologia do RNA através da associação com estruturas e sequências específicas nos alvos regulados (Hentze et al., 2018; Turner & Díaz-Muñoz, 2018). O seu papel na regulação da biologia do RNA é variável desde a transcrição, ao processamento do RNA (*capping*, *splicing*, poliadenilação e metilação), à exportação do mRNA, à montagem do ribossoma, ao início da tradução, ao alongamento e silenciamento e decaimento do RNA. Desta forma, estas proteínas conseguem regular e diversificar o transcriptoma e promover o controlo da produção e do processamento do mRNA; bem como, controlar de forma independente a expressão génica da transcrição ou da perturbação das redes transcricionais estabelecidas (Turner & Díaz-Muñoz, 2018). Todas estas funções referidas anteriormente dependem dos domínios adicionais que constituem a RBP e da posição que se liga ao RNA, sendo mediadoras ideais na regulação da expressão génica.

Portanto, o resultado do processo de expressão génica depende do equilíbrio estabelecido entre o mRNA e o processo de tradução e da sua estabilidade, sendo mediados pelas UTR's e por elementos que se ligam a essas regiões (Figura 5). Neste sentido, estudo prévios do laboratório de acolhimento pretendeu caracterizar o papel regulador dos UTR's do mRNA dos genes SMN1 e 2, cuja sequência é idêntica nestas regiões. Para tal, recorreu-se à construção de um sistema repórter heterólogo, fundindo a sequência codificante da luciferase com as regiões 5' e 3' do UTR do mRNA SMN. Os resultados deste estudo demonstraram que as duas regiões não codificantes apresentam um papel importante na regulação pós-transcriptional. Especificamente, identificaram-se três regiões reguladoras no 3'UTR do gene SMN, sendo que as extremidades apresentam um efeito negativo na regulação da expressão do mRNA, enquanto que a porção central tem um efeito positivo. Também se verificou que essa regulação ocorre ao nível da estabilidade do mRNA. Por outro lado, observou-se que a região 5'UTR do mRNA do gene SMN2 tinha um impacto negativo na tradução do gene repórter, possivelmente devido à presença de um codão de iniciação alternativo ao ATG principal (Gomes, 2010). Na sequência destes estudos, procurou-se identificar proteínas de ligação à região 3'UTR do mRNA do gene SMN2, potencialmente envolvidas na modelação da estabilidade do mRNA. Com recurso à utilização de segmentos de RNA biotinilados para captura de proteínas em extratos celulares, seguida de espectrometria de massa, foram identificadas várias RBPs, incluindo a LARP1, LARP2, LARP4 e a hnRNPG, sendo considerados candidatas à regulação pós-transcricional da expressão do SMN, (Tabela 3; (Oliveira, 2010)).

| Sonda | Extrato Celular | Proteína Identificada | Score |
|----------|-----------------|---------------------------------|-------|
| C | HEK293 | Nucleotide-binding protein-like | 136 |
| | | Histona H1 | 107 |
| | | LARP1 | 72 |
| C | SK-N-Be(2)c | Tubulina | 137 |
| | | LARP2 | 104 |
| | | hnRNP G | 58 |
| B | SK-N-Be(2)c | LARP4 | 132 |
| | | Nexina | 78 |

Tabela 3- Identificação das proteínas selecionadas nos extratos proteicos HEK293 e SK-N-Be(2)c por espectrometria de massa e respectivos “scores”. Tabela adaptada de Oliveira (2010)

1.3.3.1 Proteína LARP4 e sua família

A proteínas La Genuína, originalmente identificadas como auto-antígenos em doentes com Lupus, está presentes em todos os organismos eucariotas e são consideradas uma família proteica ancestral, encontrando-se conservadas ao longo da evolução dos seres eucariotas (Stavraka & Blagden, 2015).

A estrutura da proteínas La Genuína é caracterizada por um domínio La com uma região N-terminal composta por um motivo específico e conservado de 90 aminoácidos, designado por *La-Motif* (LAM); e ainda por um motivo de reconhecimento de RNA, designado *RNA recognition motif* (RRM) 1, separados por uma pequena ligação conhecida como domínio La ou módulo La (Stavraka & Blagden, 2015). A função da proteína La Genuína está bem caracterizada, em particular o seu papel no processamento e maturação dos transcritos da polimerase III e no processamento do pre-tRNA no núcleo. Para além disso, a La Genuína tem efeito positivo na tradução, apresentando na constituição do mRNA uma região 5'UTR longa e complexa. Porém, os seus efeitos também podem ser negativos dependendo da presença da proteína PABP e o mRNA apresenta uma

região 5'UTR curta contendo 5'TOP. (Bayfield, Yang, & Maraia, 2011; Hussain, Zawawi, & Bayfield, 2013).

O domínio LAM da proteína La Genuína é similar ao domínio LAM encontrado na família de proteínas designadas por *La Related Proteins* ou LARPs, apresentando uma conservação de 90 aminoácidos. As proteínas LARP's não possuem domínios enzimáticos, mas, como a proteína La Genuína, carregam o domínio LAM ao lado do domínio RRM formando o módulo La. Porém a localização do módulo La é diferente. No caso da proteína La Genuína, este encontra-se perto da posição N-terminal, ao contrário da família LARP que se encontra numa posição mais central. O domínio RRM é canónico nas proteínas La Genuína e LARP7, ao passo que nas restantes é uma região não típica designando-se por região *RRM_Like*. Para além disso, as proteínas La Genuína e LARP7 carregam um segundo domínio RRM – RRM2 – que não está presente nos restantes membros da família LARP (Bayfield et al., 2011; Hussain et al., 2013).

Nos humanos, encontram-se expressas cinco subfamílias da proteína LARP: a proteína La genuína (previamente conhecida como SS-B e recentemente como LARP3), LARP1 (variantes 1a e 1b/2), LARP4 (variantes LARP4 e 4b/5), LARP6 e LARP7, que podem modular a tradução do mRNA (Bayfield et al., 2011; Merret et al., 2013).

A proteína *La-related protein 4* (LARP4) é um membro da família LARP codificada, no genoma humano, no cromossoma 22, mais especificamente na região 12q13.2. O gene LARP4 é constituído por 22 exões e traduz uma proteína constituída por 724 aminoácidos, com peso molecular de 81 kDa (Bayfield et al., 2011). Não se conhece muito bem a sua função, mas sabe-se que tem um papel no metabolismo do RNA e na tradução, e que é expressa em todas as células eucarióticas (Yang et al., 2011).

Na célula, a proteína LARP4 encontra-se localizada no citoplasma associada ao polirribossoma e interage com a proteína PABPC, que se liga à região poli(A). Num estudo realizado na linha celular HeLa, os autores demonstraram que a proteína LARP4 apresenta um efeito positivo na promoção da estabilidade do mRNA (Schäffler et al., 2010; Yang et al., 2011).

Estudos anteriores do laboratório de acolhimento (Oliveira, 2010) identificaram a proteína LARP4 como tendo a capacidade de se associar à região 3'UTR do

mRNA dos genes SMN1 e SMN2 na linha celular SK-N-BE(2)c, sugerindo um papel potencial na regulação pós-transcricional da sua expressão. Esta descoberta sugere que a LARP4 pode ser um potencial candidato para modulação da estabilidade do mRNA do SMN2 para fins terapêuticos, sendo necessário para tal compreender se esta proteína exibe alguma relevância funcional em neurónios motores. Apoiando esta hipótese, Sen e seus colaboradores (2013) reportaram o gene CG11505, homólogo da LARP4 humana, como um modificador positivo do fenótipo de SMA em modelos *Drosophila melanogaster*. No entanto, tratando-se de um *screen* genético de larga escala, não foi possível inferir o mecanismo subjacente nem a natureza da interação estabelecida entre os homólogos de *Drosophila* da LARP4 e do SMN.

1.4 *Drosophila melanogaster* como modelo animal das Doenças do Neurónio Motor

A *Drosophila* são pequenas moscas, conhecidas como mosca da fruta ou do vinagre, que pertencem à ordem *Diptera* e à família *Drosophilidae*. No início do século XX, a *Drosophila* foi escolhida por Morgan e seus colaboradores, na Universidade de Columbia, para confirmar a teoria cromossômica da hereditariedade (Hales, Korey, Larracuente, & Roberts, 2015; Rubin & Lewis, 2000). Hoje em dia, a *Drosophila melanogaster* é considerada como um modelo animal de excelência, nomeadamente para o estudo das doenças do sistema nervoso, sendo utilizada em muitos laboratórios (Hales et al., 2015; Lloyd & Taylor, 2010). Porém, a sua utilização foi particularmente impulsionada quando se descobriu que o genoma Humano e da *Drosophila* apresentam um grau significativo de semelhança, a partir dos resultados obtidos nos projetos de sequenciação. Para além desta descoberta, também se verificou que 75% dos genes que se encontram associados a doença nos humanos apresentam genes homólogos em *Drosophila* (Hales et al., 2015; Lloyd & Taylor, 2010).

A *Drosophila melanogaster*, como modelo animal, apresenta um conjunto de vantagens relativamente aos modelos de animais vertebrados nomeadamente: a facilidade de distinção entre o macho e a fêmea, ter um curto ciclo de vida de cerca

de 10 dias desde o estado embrionário até ao adulto fértil e a geração de um elevado número de descendência, um tamanho reduzido e ser bastante económico em termos de manutenção. Uma das vantagens mais relevantes deste modelo animal é que permite a realização fácil de teste genéticos em grande escala. Para além disso, a anatomia da mosca encontra-se bem estudada, tendo um sistema nervoso complexo, mas que permite facilmente identificar características fenotípicas relacionadas com o desenvolvimento e a toxicidade neuronal através da observação do olho (Hales et al., 2015; Lloyd & Taylor, 2010; Nichols, 2006). Por outro lado, o modelo animal *Drosophila melanogaster* também apresenta um conjunto de desvantagens, nomeadamente o facto de se só ser possível manter as linhas vivas, não havendo possibilidade de conservação; a anatomia dos diversos sistemas ser substancialmente diferente dos humanos; e apresentar menor conservação das proteínas, entre outras (Prüßing, Voigt, & Schulz, 2013) (Tabela 4).

| Vantagens | Desvantagens |
|--|---|
| <ul style="list-style-type: none"> - Sem problemas éticos ou restrição da proteção dos animais; - Fácil e barato a manutenção em larga escala; - Fácil e barato a manipulação genética; - Curto ciclo de vida (~10 dias), curto tempo de vida (2 a 3 meses); - Sequencia do genoma completo; - Cromossomas balanceados permitem a manutenção das mutações sem genotipagem; - Mais complexo que as <i>C. elegans</i> e as leveduras; | <ul style="list-style-type: none"> - Manter apenas em culturas vivas, não permitindo a sua conservação permanente; - Pobre conservação e função das proteínas; - Anatomia dos sistemas nervoso, cardiovascular e respiratório diferente do humano; - Difícil a medição do comportamento complexo; - Não apresenta um sistema imunitário adaptativo; - Efeito das drogas no organismo pode ser significativamente diferente. |

| | |
|---|--|
| - Boa conservação das vias básicas de sinalização e dos processos celulares em geral. | |
|---|--|

Tabela 4 - Vantagens e desvantagens do modelo animal *Drosophila melanogaster*.

Tabela adaptada Prüßing et al. (2013)

1.4.1 Ciclo de vida

A *Drosophila melanogaster* é um modelo animal que tem como principal particularidade um curto ciclo de vida, que permite uma rápida geração de um grande número de descendentes para uso em estudos genéticos.

O processo de desenvolvimento, desde o ovo fertilizado até ao adulto, decorre em média entre 10 a 12 dias, à temperatura de 25°C. Porém, a temperatura influencia, no sentido inverso, o processo de desenvolvimento, visto que à temperatura de 18°C o processo de desenvolvimento ocorre em média em 19 dias (Hales et al., 2015; Nichols, 2006).

Após a fertilização, a embriogénese acontece em 24 horas, seguida de três estádios larvares (primeiro, segundo e terceiro estadio ou '*instar*'). Os primeiros dois estádios larvares decorrem, em média, em 1 dia cada um; enquanto o terceiro requer cerca de 2 dias até ao seu término. Portanto, após 5 dias da fertilização o desenvolvimento larvar está completo e inicia-se o processo de metamorfose no interior de um casulo rígido protetor, à base de quitina, formado a partir da cutícula larval externa. O processo de metamorfose depende da hormona esteroide ecdisona, que regula a expressão temporal de genes do estado larvar ao adulto. Os animais permanecem no casulo pupar durante 4-5 dias, em que se desenvolvem as estruturas que irão dar origem às estruturas adultas, a partir dos 19 discos imaginais presentes no estádios larvares. As moscas adultas eclodem e tornam-se sexualmente maduras em 8-12 horas e a partir deste momento inicia-se um novo ciclo de vida (Figura 6) (Hales et al., 2015; Nichols, 2006).

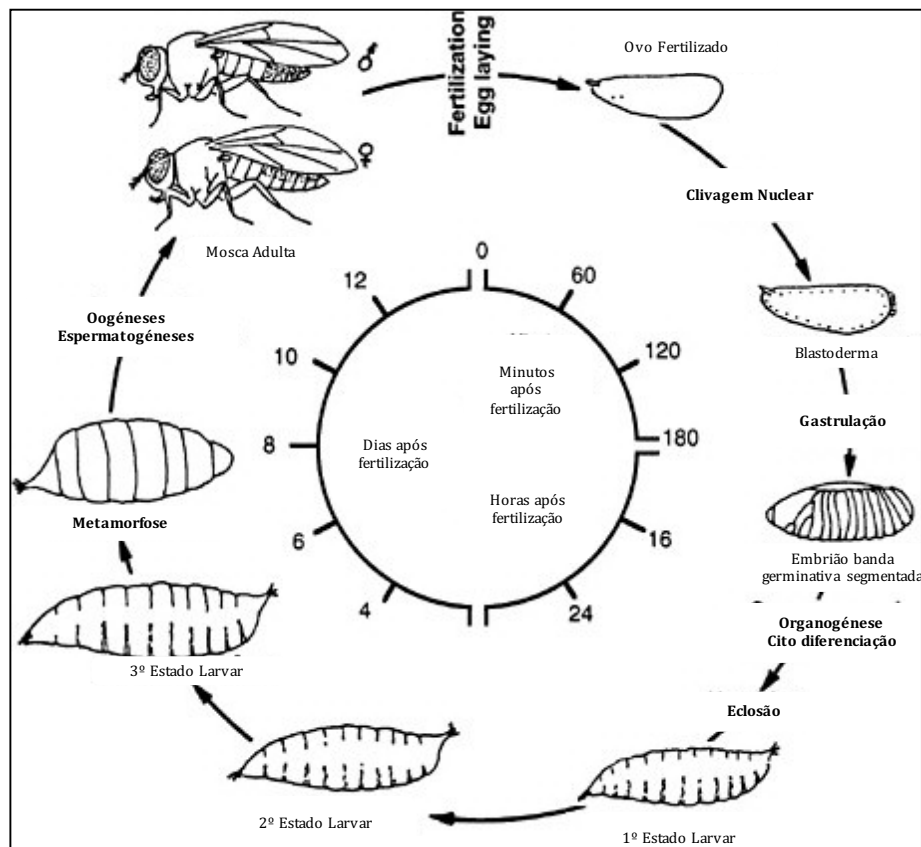


Figura 6 - Ciclo de vida da *Drosophila melanogaster*. Figura adaptada de Council (2000)

1.4.2 O sistema nervoso central da *Drosophila melanogaster*

O encéfalo da *Drosophila* contém cerca de 90.000 neurónios, um número bastante inferior ao que é observado no ser humano. Porém, ambos apresentam uma complexidade similar dos diferentes tipos de células neuronais. A comunicação entre os neurónios acontece através de contactos especializadas, que se designam por sinapses, basicamente com a mesma arquitetura proteica que no ser humano. A comunicação neuronal ocorre através da utilização dos mesmos neurotransmissores que se encontram no sistema nervoso humano, como o GABA, o glutamato e a acetilcolina, assim como as aminas biogénicas (dopamina e serotonina), e ambos os organismos possuem péptidos neuromoduladores (Venken, Simpson, & Bellen, 2011). No entanto, Venken e seus colaboradores (2011) referem que há diferenças no sistema de comunicação entre a mosca e os vertebrados. No caso do sistema de comunicação neuronal da mosca, estas usam

o glutamato como neurotransmissor nas junções neuromuscular e a acetilcolina nas sinapses sensoriais e centrais, ao contrário do que observado no sistema de comunicação neuronal dos vertebrados. Os mesmos autores referem também que as moscas possuem, como no sistema nervoso humano, canais de sódio que permitem a propagação dos potenciais de ação e a regulação do potencial de membrana através dos canais de potássio e de cálcio.

Contudo, a relação entre o número de neurónios e células da glia é diferente entre as duas espécies, apresentando o sistema nervoso da mosca, um rácio neurónio:glia de 10:1. A razão desta diferença deve-se ao facto de que nas moscas, a glia envolve um conjunto de neurónios e não um neurónio isolado (Venken et al., 2011). Esta espécie também contém muitos tipos diferentes de células da glia, como observado no ser humano.

O corpo celular dos neurónios das moscas encontra-se localizado no córtex em redor da neurópilo que é uma região composta por axónios não-mielinizados, dendrites e sinapses. Por norma, os neurónios das moscas apresentam poucas sinapses, como é o caso do sistema visual que, tipicamente, apresenta cerca de 30 a 50 sinapses por neurónio, ao contrário do que é observado no sistema visual do vertebrado que apresenta milhares de sinapses por neurónio (Venken et al., 2011).

O olho composto de *Drosophila* é uma estrutura neural composta por cerca de 750-800 unidades singulares designadas por omatídeos, que são a unidade visual funcional do olho composto. Cada omatídeo é composto por oito células fotorreceptoras, em forma de trapezoide assimétrico, e rodeadas por células cones e pigmentares primárias. A separação entre cada omatídeo é feita por doze células acessórias (Liang et al., 2016; Prüßing et al., 2013; Ready, Hanson, & Benzer, 1976).

A morte e a disfunção neuronal, assim como a perturbação das vias de desenvolvimento neuronais, cruciais à formação do olho composto, resultam em distúrbios na rede ocular, levando o fenótipo de olho irregular (*Rough Eye Phenotype* – REP). Portanto, qualquer alteração na estrutura do olho ou interrupção do desenvolvimento de uma única célula da unidade funcional resulta em fenótipos morfológicos observáveis, afetando quer o tamanho do olho, quer o tamanho das unidades funcionais e a sua pigmentação (Iyer et al., 2016). A

manifestação do fenótipo REP também pode ser induzida pela expressão de produtos genéticos tóxicos, constituindo desta forma uma leitura fácil para *screens* genéticos (Prüßing et al., 2013). A expressão de um gene de interesse pode ser direcionada para células pós-mitóticas, especificamente neurónios fotorreceptores do olho composto, com o fator de transcrição GMR-GAL4, através da utilização do sistema de expressão regulada GAL4/UAS (Kazemi-Esfarjani & Benzer, 2000) (ver abaixo).

O fenótipo de REP pode ter diferentes graus de gravidades que geralmente coincidem com diferentes níveis de degeneração das células fotorreceptoras (Prüßing et al., 2013). Desta forma, alterações no declínio neuronal podem ser observadas com recurso a microscopia óptica (Prüßing et al., 2013). Porém, esta degeneração do olho não pode imitar completamente os processos complexos que levam à doença no cérebro humano.

1.4.3 Genoma e ferramentas de manipulação genética

O genoma de *Drosophila* apresenta um tamanho relativamente reduzido, tendo sido completamente sequenciado e anotado (Adams et al., 2000). É composto por cerca de 13.600 genes, em comparação com os cerca de 25.000 genes do genoma humano. Para além da diferença que se observa no número de genes entre os dois organismos, a sequenciação do genoma de *Drosophila* verificou que mais de 90% dos seus genes são semelhantes a genes presentes nos genomas humano e de ratinho (Adams et al., 2000).

O genoma de *Drosophila* é composto por 4 cromossomas, sendo denominados por X/Y, 2, 3 e 4 (Figura 7). O primeiro cromossoma corresponde ao cromossoma sexual. O cromossoma sexual pode ser o acrocêntrico X ou o cromossoma submetacêntrico Y, que é composto maioritariamente por heterocromatina. Uma particularidade do cromossoma Y é a presença de alguns genes que são importantes para a espermatogénese, mas que não são importantes para a sobrevivência do animal. Os restantes três pares de cromossomas são autossomas e contêm a maioria dos genes. Os cromossomas 2 e 3 são autossomas metacêntricos com os braços de cada lado do centrómero designados por L e R.

Comparativamente com os outros, o cromossoma 4 é muito pequeno.

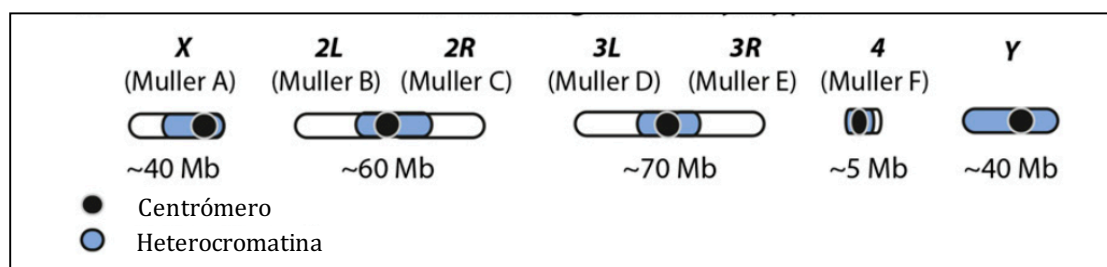


Figura 7 - Esquema da organização dos cromossomas de *Drosophila*. Imagem adaptada de (Hales et al., 2015)

A *Drosophila* apresenta uma grande vantagem do ponto de vista de manipulação genética que é a existência de cromossomas balanceados (St Johnston, 2002). Estes cromossomas resultam de rearranjos genéticos resultando em múltiplas inversões cromossômicas aninhadas que impedem a recombinação homóloga. Estes cromossomas apresentam genes marcadores que facilitam a sua identificação e permitem a manutenção de alelos mutantes recessivos que de outra forma seriam letais. Assim, os cromossomas balanceados são capazes de suprimir o *crossing over* com o cromossomas homólogo durante a meiose e, mesmo que este ocorra, os cromatídeos resultantes carecem de alguns genes ou possuem genes duplicados que impede a sobrevivência da descendência (St Johnston, 2002). Como referido anteriormente, os cromossomas balanceados são marcados com fenótipo visíveis, através de uma ou mais mutações dominantes, que permite seguir o balanceamento no decorrer dos cruzamentos. Portanto, a principal vantagem do recurso ao uso dos cromossomas balanceados é a capacidade de manutenção do alelo mutante (Casso, Ramírez-Weber, & Kornberg, 1999; St Johnston, 2002).

Outra vantagem do modelo de *Drosophila* é o uso de sistemas de expressão binários de origem microbiana, constituídos por dois elementos genéticos: um fator de transcrição exógeno e um promotor sintético que contém a sequência de ligação do fator de transcrição (Okray & Hassan, 2013). Ao longo dos anos foram desenvolvidas diversas linhas genéticas de *Drosophila* que apresentam o fator de transcrição exógeno sob controlo de uma região promotora específica que conduz à sua expressão. Quando expresso, este fator de transcrição irá ligar-se, especificamente, ao promotor sintético e

dirigir a expressão do alvo de interesse, seja um gene específico, um marcador ou um RNA de interferência (RNAi). O principal sistema de expressão gênica regulada utilizado em *Drosophila* é o sistema GAL4/UAS, para além dos sistemas LexAop e QF:UAS, que podem ser utilizados em paralelo com o sistema GAL4/UAS (Brand & Perrimon, 1993; Okray & Hassan, 2013; Potter et al., 2010; Szüts & Bienz, 2000).

1.4.3.1 GAL4/UAS como um sistema de expressão gênica binário

A regulação da expressão gênica pode ser direcionada por elementos reguladores, em tempo e local específicos, se colocados a montante da sequência desejada (Brand & Perrimon, 1993; Venken et al., 2011). De acordo com Venken e seus colaboradores (2011), um sistema de regulação da expressão gênica binário é preferencial, visto apresentar um conjunto de vantagens como a expressão restrita de proteínas tóxicas, a amplificação dos níveis de proteína e a capacidade de expressão de diferentes repórteres e efetores em células específicas ou a mesma resposta em diferentes tipos celulares.

O sistema binário consiste na tradução de um fator de transcrição, que pode ser expresso em diferentes células ou em diferentes tempos do desenvolvimento, que irá ter efeito num local específico e, por conseguinte, no processo de transcrição do gene de interesse a jusante da região de ligação do fator de transcrição (Brand & Perrimon, 1993; Venken et al., 2011). Porém, o fator de transcrição pode ser reprimido por repressores que atuam de forma direta sobre ele ou por compostos que ativam ou inativam o fator de transcrição ou o repressor, permitindo o controlo temporal e espacial da expressão genética (Venken et al., 2011).

No modelo animal de *Drosophila*, o sistema binário GAL4/UAS foi o primeiro a ser desenvolvido e é o mais utilizado (Fischer, Giniger, Maniatis, & Ptashne, 1988). O GAL4 é um fator de transcrição, derivado de leveduras, que se liga a uma sequência específica designada por *Upstream Activating Sequences* (UAS), permitindo o início da transcrição do gene de interesse que se encontra a jusante (Venken et al., 2011) (Figura 8). Porém, a sua atividade pode ser inibida pelo repressor GAL80.

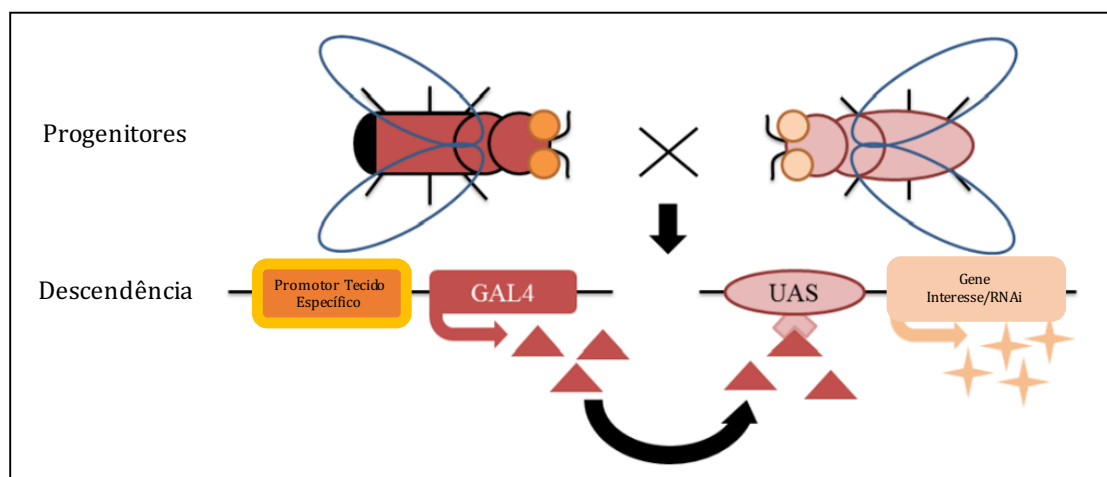


Figura 8 - Esquema do sistema de expressão binário GAL4/UAS. Imagem adaptada de Tan & Azzam (2017)

O sistema binário, como modelo de regulação de expressão gênica em *Drosophila*, permite uma rápida geração de animais em que a expressão do gene de interesse pode ser direcionada para diferentes tipos de células e tecidos. Para além desta característica, o fator de transcrição (GAL4) e o gene de interesse encontram-se separados em duas linhas de animais transgênicos distintos, sendo ambos os progenitores viáveis. Isto significa que uma linha transgênica contém o gene de interesse, como o UAS-GFP, na ausência do fator de transcrição e a outra linha transgênica contém a proteína ativadora sem o gene de interesse. Após o cruzamento das duas linhas, ocorre a ativação do gene de interesse e, desta forma, a expressão do fenótipo, da sua ativação, nos progenitores. Portanto, os descendentes deste cruzamento expressam quer o fator de transcrição, quer o gene de interesse (Brand & Perrimon, 1993).

Porém, o sistema binário tradicional GAL4/UAS permite apenas o controle regional da expressão do gene. Desta forma, seria importante o desenvolvimento de metodologias que permitissem o controle temporal da expressão gênica e que assim fosse possível o estudo da expressão de um gene em momentos críticos e em tecidos específicos para o desenvolvimento de um fenótipo em particular.

Uma das abordagens que permite a regulação temporal da expressão gênica será através do recurso à proteína de levedura GAL80 (Lue, Chasman, Buchman, & Kornberg, 1987). A proteína GAL80 irá reprimir a atividade da proteína GAL4 através da sua ligação à região C-terminal da proteína GAL4. Desta forma, a GAL80

impede que a GAL4 se ligue à sequência UAS e, por conseguinte, não permite a ativação da transcrição do gene de interesse (Lai & Lee, 2006). Porém, esta abordagem permite apenas o controlo da expressão génica de forma temporal. Por isso, foi desenvolvido uma nova abordagem que permite o controlo quer temporal quer espacial da regulação da expressão génica, com recurso ao sistema GAL4/UAS, através do uso da temperatura, denominada por técnica TARGET (McGuire, Le, Osborn, Matsumoto, & Davis, 2003). Estes autores desenvolveram um sistema em que clonaram uma versão da GAL80 sensíveis à temperatura, GAL80^{TS}. A sequência desta proteína apresenta uma alteração na posição 203, onde ocorre uma substituição do aminoácido Glicina por um aminoácido de Arginina. A proteína GAL80^{TS} tem uma expressão ubíqua a uma temperatura de 19°C, impedindo a atividade da GAL4. Quando há uma mudança de temperatura para os 30°C, a atividade da GAL4 não é impedida pela GAL80^{TS} (Nicholson et al., 2008). Portanto, o controlo temporal do gene de interesse, dependente da GAL4, pode ser controlada através de alterações da temperatura do sistema devido à repressão da proteína GAL4 pela GAL80^{TS} (Figura 9A). Desta forma, a técnica TARGET apresenta uma vantagem relativamente à técnica da GAL80, visto a GAL80^{TS} ser expressa de forma ubíqua e estar sob o controlo do promotor da tubulina 1 α , e não do *heat shock*, e dessa forma ser apenas ativada a uma temperatura permissiva, permitindo dessa forma o controlo da transcrição do gene de interesse (McGuire et al., 2003). O sistema TARGET é totalmente compatível com a vasta gama de linhas GAL4 que se encontram disponíveis, sendo amplamente utilizado no estudo do desenvolvimento e da função neuronal em *Drosophila* (Nicholson et al., 2008).

Outra abordagem de regulação temporal da expressão genética foi desenvolvida, através da utilização do sistema GAL4/UAS, em que as variantes de GAL4 são induzidas por hormonas (Figura 9B). Estas podem ser de dois tipos: GAL4-receptor de estrogénio (GAL4-ER) (Han, Stein, & Stevens, 2000) e *GeneSwitch* (Osterwalder, Yoon, White, & Keshishian, 2001; Roman, Endo, Zong, & Davis, 2001). Esta abordagem necessita de uma proteína em que há a combinação do domínio de ligação ao DNA da GAL4 com o domínio de ativação transcriptional do recetor de uma hormona esteroide, para que haja ligação do ligante e, dessa forma, se tornar transcricionalmente ativo. No caso da variante GAL4-ER, esta resulta da

fusão da GAL4 com o domínio de ligação do recetor humano de estrogénio, sendo ativada pelo estradiol (Han et al., 2000). Por outro lado, a variação *GeneSwitch* é o resultado da fusão da GAL4, do domínio de ligação do recetor da progesterona humano e parte de ativação do domínio da subunidade p65 do NF-kappa-B humano. Esta variação é ativada pela ligação do esteroide sintético Mifepristona (RU-486) (Osterwalder et al., 2001; Roman et al., 2001). A oferta das hormonas esteroide pode ser feita de duas formas: através da alimentação ou através da imersão dos animais em uma solução de esteroide. Contudo, a expressão é diferente consoante a variação utilizada no controlo da expressão génica, sendo a variação GAL4-ER mais lenta do que o *GeneSwitch*. A variação *GeneSwitch* permite que a deteção da expressão seja possível após 3 a 5 horas da exposição, apresentando uma expressão máxima às 21-48 horas (Han et al., 2000; Osterwalder et al., 2001; Roman et al., 2001). Porém, os níveis de expressão podem ser diferentes consoante a dosagem de esteroide administrado, dependendo do efeito de dose-resposta (Osterwalder et al., 2001).

Para que seja necessário a utilização deste sistema binário de expressão, é necessário a criação de um vetor de expressão que contenha o gene de interesse. Desta forma é necessário a clonagem do gene de interesse no local de clonagem do vetor de expressão pP(UAST), que se encontra a jusante do local de ligação ao GAL4 e da caixa sintética TATA e a montando do SV40 e do sítio de poliadenilação. Posteriormente, o vetor de expressão é injetado nos embriões e o efeito nas moscas transgénicas depende das diferentes inserções, verificando que diferentes inserções são expressas a diferentes níveis (Phelps & Brand, 1998).

A expressão do gene de interesse resulta do cruzamento entre o vetor de expressão e a linha *driver* GAL4, sendo possível avaliar o fenótipo na descendência. Após o cruzamento, o primeiro passo da análise será determinar onde o gene de interesse está a ser expresso. A expressão ectópica pode ser realizada diretamente através da marcação com anticorpos contra o produto do gene de interesse ou através de hibridação *in situ* para o mRNA do gene de interesse. Porém, será necessário distinguir entre o produto do gene endógeno do gene que se encontra fundido com o UAS. Por isso, será uma vantagem a fusão da região codificante do gene de interesse com uma molécula tag (Phelps & Brand, 1998).

E, por fim, a determinação das características espaço-temporal do gene de interesse será possível perceber as consequências biológicas da sua expressão.

O sistema binário GAL4/UAS também pode ser utilizado para analisar a função do gene de interesse quando há diminuição dos seus níveis de expressão. Desta forma, foi desenvolvida uma metodologia de RNAi que permite a redução ou a eliminação do mRNA do gene de interesse e dessa estudar a sua função quer em animais transgênicos quer em linhas celulares (Kennerdell & Carthew, 2000).

O mecanismo de RNAi é um mecanismo celular endógeno que é desencadeado por uma molécula de RNA de cadeia dupla (dsRNA), resultando na degradação de RNAs homólogos, sendo primariamente descrito em *C. elegans*. Em *Drosophila*, este mecanismo pode ser induzido através da microinjeção, na blastoderme do embrião, de um dsRNA ou de uma construção que contenha o RNAi do gene de interesse (UAS-RNAi) (Heigwer, Port, & Boutros, 2018).

A injeção de dsRNA (entre 200-2000bps) ou siRNAs (21-22 nucleótidos) de um gene de interesse, em embriões de *Drosophila*, foi o primeiro procedimento a ser utilizado para induzir a diminuição da expressão desse gene. Porém esta técnica apenas poderia estudar os efeitos durante o desenvolvimento embrionário (Heigwer et al., 2018; Kennerdell & Carthew, 2000; Perrimon et al., 2013).

Para colmatar este problema foi desenvolvido a abordagem transgênica em que se recorre ao sistema binário GAL4/UAS para expressar RNAs curtos, com repetição invertida, em forma de grampo com a função de atingir um gene específico no sentido de reduzir ou eliminar a expressão do mRNA (Kennerdell & Carthew, 2000). Porém, estas alterações não são permanentes no transcrito, apenas reduzem os níveis de mRNA do gene de interesse (Hales et al., 2015).

Inicialmente esta abordagem apresentava algumas desvantagens visto utilizarem inserções aleatórias dos transgênicos, que continham as construções de interferência especificamente para um gene de interesse. Contudo, esta abordagem era ineficaz na capacidade de redução da expressão do RNA para níveis nulos e poderiam mesmo produzir letalidade ou esterilidade, dependendo do local de inserção (Hales et al., 2015).

Para colmatar este problema, foi desenvolvido um sistema de integração phiC31a/AttP, no sentido de integrar a construção UAS-RNAi do gene específico

numa região do cromossoma que garantisse altos níveis de expressão e que não prejudicasse outro gene (Kennerdell & Carthew, 2000; Ni et al., 2008).

Atualmente, no *Drosophila* RNAi Screening Center and Vienna *Drosophila* Research Center encontram-se disponíveis um conjunto de bibliotecas que permite realizar RNAi em cerca de 90% dos genes que codificam proteínas. Para além disso, a existência de linhas GAL4 bem caracterizadas e que são ativadas em locais e estádios de desenvolvimento específicos, permite o uso de RNAi para melhor caracterizar a função dos genes nos processos celulares e de desenvolvimento.

Neste sentido, através da criação de um vetor de expressão que contenha o gene de interesse ou sequências repetidas invertidas pode ser possível analisar a função desse mesmo gene, de forma temporo-espacialmente, com recurso ao sistema de expressão GAL4/UAS.

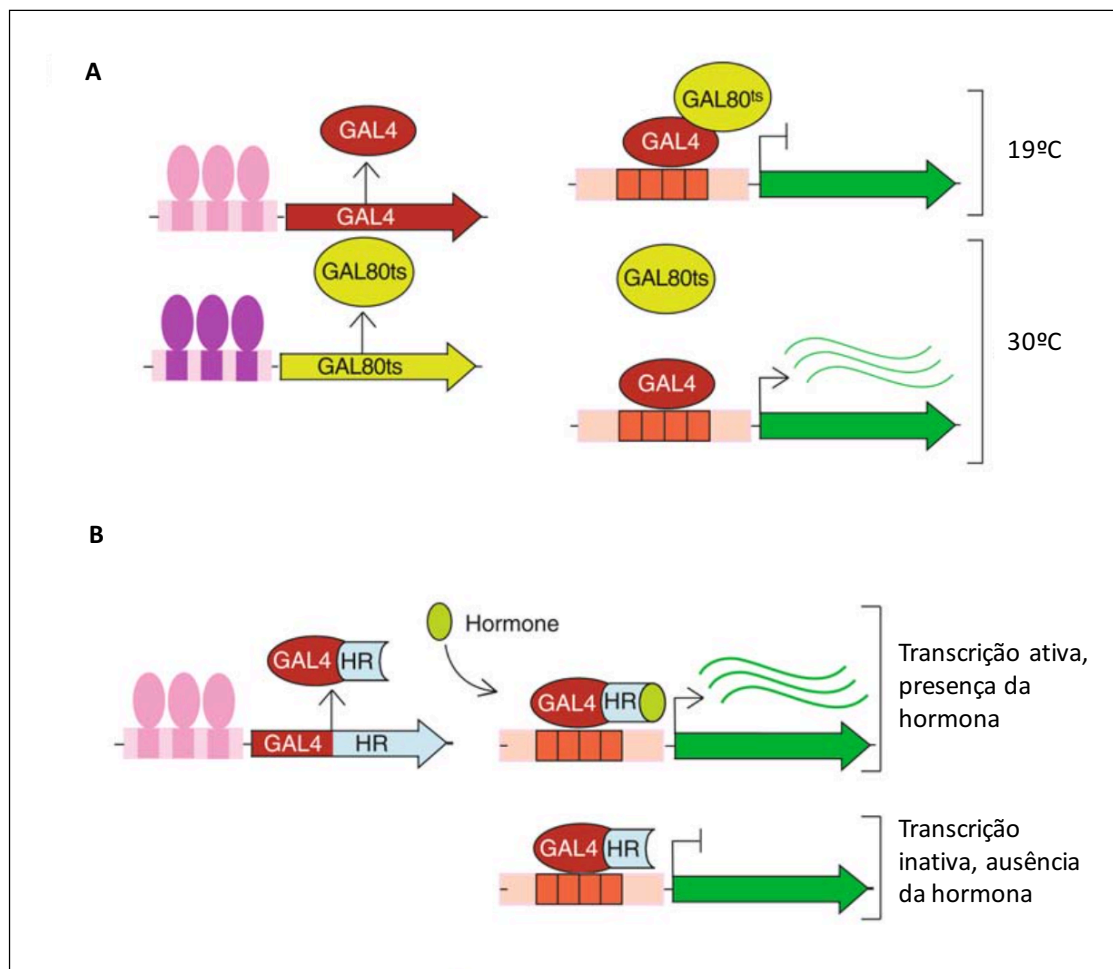


Figura 9 - Esquemas de modificações do sistema de expressão binário GAL4/UAS: Sistema TARGET (A) e Sistema GAL-ER e GeneSwitch (B). Imagem adaptada de Tan & Azzam (2017)

II. Objetivos

O laboratório de acolhimento tem explorado os mecanismos de regulação da expressão do mRNA dos genes SMN1 e 2 como potenciais alvos terapêuticos para a SMA. Trabalhos realizados previamente revelaram a existência de mecanismos regulatórios negativos e positivos envolvendo os UTRs do mRNA SMN1/2, e identificaram um conjunto de proteínas como potenciais modeladores destes eventos, influenciando os níveis celulares da proteína SMN (Gomes, 2010; Oliveira, 2010). Em particular, Oliveira (2010) identificou diversos membros da família LARP, incluindo a proteína LARP4, como potenciais reguladores da região 3'UTR do mRNA do SMN1/2.

Com estas observações, abre-se uma nova perspectiva na biologia da SMA, uma vez que não é conhecida a função desta família de proteínas ao nível do sistema nervoso e, mais especificamente, a sua ligação à patologia SMA. Evidências em suporte da existência desta ligação foram obtidas em *screens* de larga escala em *Drosophila melanogaster* (Chang et al., 2008; A. Sen et al., 2013).

O objetivo geral deste trabalho é, assim, o de caracterizar o papel da proteína LARP4 no sistema nervoso central e os seus efeitos na expressão do SMN e no fenótipo da SMA.

Para a sua concretização, foram definidos os seguintes objetivos específicos:

- 1) Caracterização da expressão da proteína LARP4 nas linhas celulares humanas HEK293, SK-N-BE(2)c e GM03813 e em embriões de ratinho, com foco particular nas junções neuromusculares;
- 2) Caracterização da interação funcional entre a proteína LARP4 e a expressão do gene SMN em linhas celulares humanas (HEK293WT e SK-N-BE(2)c), nas condições de sobre- e sub-expressão;
- 3) Estudo comparativo e clonagem do ortólogo da LARP4 de *Drosophila melanogaster* (CG11505) no vetor de expressão pUAST, como ferramenta para estudar a função desta proteína em *Drosophila melanogaster*;
- 4) Caracterizar o impacto da sobre e sub-expressão de CG11505 no sistema nervoso em *Drosophila melanogaster*.

III. Materiais e Métodos

3.1 Estudo da expressão e função da LARP4

3.1.1 Linhas celulares e cultura de células

A linha celular HEK293 foi mantida em meio *Minimum Essential Medium + Glutamax-I* (Gibco®) suplementado com 10% (v/v) *Fetal Bovine Serum* (FBS) e 1% (v/v) *Non-Essential Amino Acids* (Gibco®) à temperatura 37°C e 5% CO₂.

A linha celular SK-N-Be(2)c foi mantida em meio 45% (v/v) *Dulbecco's Modified Eagle Medium* (Gibco®), 45% (v/v) *F12 medium* (Gibco®) e suplementado com 10% FBS à temperatura de 37°C e 5% CO₂.

A linha celular GM03813 (Coriell Institute for Medical Research®) foi mantida em meio *Minimum Essential Medium + Glutamax-I* (Gibco®) suplementado com 10% (v/v) FBS e 1% (v/v) *Non-Essential Amino Acids* (Gibco®) à temperatura 37°C e 5% CO₂.

Cada linha celular foi sub-cultivada a cada 2-3 dias.

3.1.2 Transfecção celular

As linhas celulares HEK293 e SK-N-Be(2)c foram transfectadas para estudar o efeito da sobre- ou sub-expressão do gene LARP4.

Para estudar o efeito do aumento da expressão da LARP4, as linhas celulares HEK293 e SK-N-Be(2)c foram transfectadas com os plasmídeos GFP (utilizado como controlo) e GFP-LARP4 com o GeneJuice (Merck®), seguindo as instruções do fabricante, em placas de 24 poços. No caso da linha celular HEK293, 1,25x10⁵ células em cada poço foram transfectadas com 250ng de cada plasmídeo, durante 24 horas. No caso da linha celular SK-N-Be(2)c, foram transfectadas 6x10⁴ células em cada poço com 500ng de cada plasmídeo (GFP e GFP-LARP4), durante 72 horas. Após o período de transfecção, as células foram observadas ao microscópio de fluorescência (Leica DMI6000B) no sentido de analisar a eficiência de transfecção.

Para analisar o efeito da diminuição da expressão da LARP4, recorreu-se à transfecção de RNA de interferência para a inativação do gene LARP4, com auxílio

do agente de transfecção siPORT™ NeoFX™ (ThermoFisher Scientific®), seguindo as instruções do fabricante. Para todas as linhas celulares, o período de incubação foi de 72 horas e a experiência foi realizada em placa de 24 poços. Para as linhas celulares HEK293 e SK-N-Be(2)c foram transfectadas 5x10⁴ células de cada linha celular com 20pmol de siRNA controlo negativo (*Silencer*™ Select Negative Control No. 1, ThermoFisher Scientific®) e 20pmol de siRNA LARP4 (s41455, ThermoFisher Scientific®). A eficiência de inativação do gene LARP4 foi avaliada através de reacção de PCR em tempo real.

3.1.3 Ensaio de Luciferase

A linha celular HEK293 foi utilizado nos ensaios de luciferase para analisar os efeitos da sobre e sub-expressão da LARP4 sob o gene repórter SMN-Luciferase. Para estudar o efeito de sobre-expressão, a linha celular HEK293 foi transfectada com os plasmídeos GFP e GFP-LARP4 e co-transfectadas com os plasmídeos Luciferase-3'UTR SMN1/2 e pcDNA3-renilla, durante 24 horas. Para o processo de transfecção, foi utilizado o reagente GeneJuice (Merck®), seguindo as instruções do fabricante, transfectando 200ng dos plasmídeos GFP ou GFP-LARP4 e 25ng de Luciferase-3'UTR SMN1/2 e pcDNA3-renilla, em placa de 24 poços.

No caso do estudo da sub-expressão da LARP4, a mesma linha celular foi transfectada com os RNA de interferência, com recurso ao agente de transfecção siPORT™ NeoFX™ (ThermoFisher Scientific®), seguindo as instruções do fabricante, por um período de 72 horas. Após 24 horas da transfecção, as células foram transfectadas com os plasmídeos Luciferase-3'UTR SMN1/2 e pcDNA3-renilla, com recurso ao reagente GeneJuice (Merck®), seguindo as instruções do fabricante. A análise da sub-expressão da LARP4 foi realizada segundo a técnica de PCR em tempo real, como explicitada posteriormente na secção 3.1.5.

A avaliação dos níveis de luciferase foi com recurso ao sistema *Dual-Luciferase*® Reporter (Promega®), de acordo com as indicações do fabricante. Na análise de dados do ensaio utilizou-se como controlo interno o pcDNA3-renilla.

3.1.4 Isolamento e extração de RNA

O RNA total das células transfetadas foi isolado através da utilização do reagente *TRI Reagent Solution*® (ZymoResearch®), de acordo com as instruções do fabricante. O RNA foi eluído em DPEC-H2O e guardado a -80°C. A concentração do RNA foi quantificada no Nanodrop Spectrophotometer ND-1000.

3.1.5 Transcrição reversa e PCR em tempo real

O cDNA foi sintetizado a partir de 500ng de RNA total com recurso *M-MuLV Transcriptase Reverse* (NZYTech®) e Oligo d(t) (NZYTech®), segundo as instruções do fabricante. Previamente à reação de transcrição reversa, as amostras sofreram uma reação com a enzima DNase I (ThermoFisher®) durante 30 minutos à temperatura de 37°C, com posterior inativação durante 20 minutos à temperatura de 70°C.

A análise da expressão génica foi realizada através de reação de PCR em tempo real, utilizando como produto de reação o SsoFast™ EvaGreen® Supermix (Bio-Rad®) e seguindo as instruções do fabricante, no aparelho CFX96 (Bio-Rad®). As condições de reação foram: 1) amplificação: 95°C, 30sec; [95°C, 5sec; 60°C, 5sec]x40 ciclos; 2) Curva de *Melting*: entre 65°C e 95°C, durante 5 segundos, com aumentos de 0,5°C. Os primers utilizados para a amplificação encontram-se descritos na tabela 5. O gene GAPDH foi usado como gene de referência em todas as amostras. A análise de expressão génica foi realizada de acordo com o método $\Delta\Delta C_T$.

| Primer | | Sequência |
|-------------------------|-----|----------------------------------|
| <i>Hs_SMN</i> | FW | 5'- TGCTGGCTGCCTCCATTT-3' |
| | Rev | 5'- GCATCATCAAGAGAATCTGGACAT -3' |
| <i>LARP4_all</i> | Fw | 5'- AGCTACATCAGGTGCTCATCC - 3' |
| | Rev | 5'- GCTTTCTCCGGAAACATCAT- 3' |
| <i>GAPDH</i> | Fw | 5' - TGCACCACCAACTGCTTAGC - 3' |
| | Rev | 5' - GGCATGGACTGTGGTCATGAG - 3' |

Tabela 5 - Lista de primers utilizados nos ensaios de PCR em tempo real

3.1.6 Extratos proteicos

As células foram lavadas uma vez com PBS1x frio e posteriormente lisadas com 2xSDS-page sample buffer (1M Tris-HCl pH 6.8, 10% SDS, Glicerol, β -mercaptoetanol, 1% Azul de Bromofenol), Benzonase (1:100, Sigma[®]) e 0,5M MgCl₂ (NzyTech[®]), adicionada diretamente no poço com 10 minutos de incubação à temperatura ambiente (TA). Posteriormente foram desnaturadas à temperatura de 98°C, durante 5 minutos e guardadas a -20°C.

3.1.7 Western-Blot

10 μ L de proteína de cada amostra foi separada, por electroforese, num gel de 10% de SDS-poliacrilamida em *Tris/Glycine/SDS running buffer* (Bio-Rad[®]). Após o processo de separação, as proteínas foram transferidas para membrana de PVDF (Millipore[®]), em meio frio *Tris/Glycine transfer buffer* (Bio-Rad[®]), durante 1 hora a 100V. A membrana de PVDF foi bloqueada com uma solução de bloqueio constituída por 5% (m/v) de leite magro em PBS-0,1% (v/v) Tween20 (PBS-T),

durante 1 hora, em agitação. Os anticorpos primários e secundários foram preparados na solução de bloqueio. Os anticorpos primários anti-SMN (1:1000, #610646, BD Bioscience[®]) e anti- α -tubulina (1:3000, 05-829, Merck Millipore[®]) foram incubados durante a noite em agitação, a 4°C. No dia seguinte, foram realizadas quatro lavagens da membrana com PBS-T, cada uma durante 10 minutos, e a membrana incubada com o anticorpo secundário *anti-mouse*-HRP (1:4000, #1721011, Bio-Rad[®]) durante 1 hora, à TA, em agitação. O anticorpo secundário não ligado foi eliminado por lavagem, de forma idêntica ao descrito anteriormente. No final, as proteínas foram visualizadas com recurso ao reagente Clarity[™] Western ECL Blotting, no aparelho ChemicDoc (BioRad[®]). Na análise da expressão proteica, a proteína α -tubulina foi utilizada com referência.

3.1.8 Imunocitoquímica

As células de cada linha celular cresceram em lamelas 10X10mm em P35 durante 24 horas. Após esse período, as células foram lavadas com PBS 1x e foram fixadas com uma solução de 3,4% (v/v) de Formaldeído em PBS 1x, durante 10 minutos, à TA. Após o processo de fixação, as células foram novamente lavadas com PBS 1x durante 5 minutos, em três períodos, permeabilizadas com 0,5% (v/v) Triton X-100 em PBS 1x, durante 10 minutos à TA, e novamente lavadas com PBS 1x. As células foram coradas com os anticorpos primários anti-SMN (1:50, sc-32313, Santa Cruz Biotechnology[®]) e/ou anti-LARP4 (1:50, HPA039306, Sigma[®]) por um período de 1 hora, em ambiente húmido. Posteriormente, sofreram um processo de três lavagens de PBS-T, durante 5 minutos, e incubadas com os anticorpos secundários *anti-mouse* Alexa 568 (1:500, A10037, Invitrogen[®]) e *anti-rabbit* Alexa 488 (1:500, A21206, Invitrogen[®]) durante 1 hora, em ambiente húmido e escuro. Os anticorpos primários e secundários foram diluídos numa solução de PBS-T.

As células foram observadas no microscópio de fluorescência Olympus BX 60.

3.1.9 Imunohistoquímica

A técnica de Imunohistoquímica foi realizada em embriões de ratinho E13,5 e E15,5, em cortes no plano sagital com uma espessura de 12 µm no criostato à temperatura de -25°C, secos na lâmina durante 45 minutos à TA. Após este período, os cortes foram selados com *DAKO Pen* e submetidos a três lavagens, de 5 minutos cada, com PBS 1x. A permeabilização foi realizada com a solução de Triton-X-100 0,2% (v/v) em PBS 1x durante de 20 minutos, seguido de um processo de lavagem similar ao anterior. Após a lavagem, o bloqueio foi efetuado através do uso do *Mouse on Mouse (M.O.M.TM) Basic Kit* (Vector Laboratories[®]), segundo instruções do fabricante. Os cortes histológicos foram incubados durante a noite à temperatura de 4°C, com os anticorpos primários diluídos no *Protein Concentrator* do kit: anti-LARP4 (1:50, HPA039306, Sigma[®]), anti-Neurofilamento 160 (1:100, N5264, Sigma[®]), α -Bungarotoxina Alexa FluorTM 488 (1:200, B13422, Invitrogen[®]) e anti-sinaptofisina (1:50, M7315, Agilent[®]). No dia seguinte, realizou-se um conjunto de três lavagens de 15 minutos cada com PBS1x, e os cortes histológicos foram incubados com os anticorpos secundários correspondentes à TA durante uma hora e trinta minutos. Novamente, foi realizado o processo de lavagem, durante 4x5 minutos, com PBS. Por fim, os cortes foram corados com o DAPI durante 30 segundos e realizada a sua montagem com lamela e meio de montagem de propilgalato. Os cortes histológicos foram observados no microscópio de fluorescência Olympus BX 60.

3.1.10 Análise Estatística

A análise estatística foi realizada através do teste estatístico t de Student para duas amostras emparelhadas. A análise e a apresentação gráfica realizou-se com recurso ao software GraphPad Prism Versão 6 (GraphPad Software[®]). Os Resultados são apresentados como médias \pm desvio padrão.

3.2 Estudo da função da proteína CG11505

3.2.1 Restrição enzimática, vetor e insert

O plasmídeo pUAST_CG11505 foi construído através da inserção do produto de PCR do gene CG11505 no vetor pUASTattb (oferta de Aaron Voigt). Para a obtenção do produto de PCR, foi usado como molde o plasmídeo pDNR_Dual_CG11505 (AT2581, oferta de Aaron Voigt) e a enzima NZyTaq Polimerase (NZYTech[®]), com os primers referenciados na tabela 6, por forma a introduzir um “tag” N-terminal de hemaglutinina. As condições da reação de PCR foram: 95°C, 2min; [95°C, 1min; 55°C, 1min; 72°C, 3min]x10 ciclos; [95°C, 1min; 65°C, 1min; 72°C, 3min]x25 ciclos; 72°C, 5min. Após o PCR, o plasmídeo e o produto de PCR foram digeridos com recurso às enzimas de restrição NotI (ThermoFisher Scientific[®]) e XbaI (ThermoFisher Scientific[®]) durante uma hora à temperatura de 37°C, seguida de um período de inativação durante vinte minutos a 80°C. A eficiência da reação de digestão foi verificada através de eletroforese em gel de agarose 0,8% durante uma hora, confirmando os tamanhos das bandas obtidas.

| Primer | Sequência | Enzima Restrição |
|------------|---|------------------|
| Fw | ATGCGGCCGCATGTACCCATACGATGTTCCAGATTACGCTATGAACGGAGATGCACT | NotI |
| Rev | GGCTCGAGTCAGAAGAGGCCGCGGCCTTGTGGT | XbaI |

Tabela 6 - Lista de primers utilizados no PCR convencional no processo de clonagem do vetor pUAST-CG11505

Após confirmação da eficiência da reação de digestão, e do correto tamanho das bandas obtidas, as mesmas foram extraídas do gel de agarose e purificadas com recuso ao kit NZYGelpure (NZYTech[®]), seguindo as informações do fabricante. As

amostras foram eluídas em água Milli Q e quantificadas no Nanodrop Spectrophotometer ND-1000.

3.2.2 Reação de ligação

A reação de ligação foi realizada com proporções vetor:insert de 1:0 (controle – sem insert) e 1:3, sendo a quantidade de vetor utilizada de 30ng. A reação de ligação foi realizada com recurso à enzima T4 DNA Ligase (NZYTech®) durante 16 horas à temperatura de 20°C.

3.2.3 Transformação bacteriana

A transformação bacteriana foi realizada com recurso a células competentes NZY5a (NZYTech®), previamente descongeladas em gelo. A transformação resultou da adição de 5µL da reação de ligação às células competentes, seguidas de um período de incubação, em gelo, durante 30 minutos. Após este período, foi realizado um choque térmico a 42°C durante 45 segundos e, novamente colocadas em gelo durante 2 minutos. A cada reação de transformação foram adicionados 900µL de meio líquido *Luria Broth* (LB) [1% w/v Tryptone; 0.5% w/v Yeast extract; 171 mM NaCl] sem antibiótico, e as células incubadas em agitação (225rpm) durante uma hora a 37°C. Terminado o período de incubação, as suspensões bacterianas foram semeadas em placas de meio LB Agar [15g/L Agar, 10g/L Tryptone, 5g/L Yeast extract e 10g/L NaCl] contendo 100 µg/mL de Ampicilina, e incubadas durante a noite a 37°C.

3.2.4 Extração e purificação de plasmídeos (Miniprep)

As colónias foram inoculadas em 4mL de meio líquido LB com 100µg/mL de Ampicilina durante 12 a 16 horas, em agitação (225rpm), à temperatura de 37°C. Após o período de incubação, foi realizada a extração e purificação do plasmídeo

com recurso ao kit NZYMiniprep (NZYTech[®]). Inicialmente, foi realizada a centrifugação da mistura a 15.000rpm, à TA, durante 30 segundos, e retirado o sobrenadante. O *pellet* foi re-suspendido em 250uL do Buffer A1. Posteriormente, foram adicionados 250uL do Buffer A2 e a suspensão foi misturada gentilmente por inversão, durante 8 vezes, e incubada à TA durante 4 minutos. À solução de lise celular foram adicionados 300uL de Buffer A3, misturados gentilmente, por inversão, durante 6 vezes. O lisado celular foi centrifugado a 15.000rpm durante 10 minutos à TA. O sobrenadante foi colhido e colocado em colunas de centrifugação NZYTech[®], centrifugadas à velocidade 11.000g durante 1 minuto à TA. Foram adicionados à coluna 500uL de Buffer AY, e a coluna novamente centrifugada durante um minuto a 15.000rpm à TA. Por fim, foram adicionados 600uL do Buffer A4 à coluna, que se centrifugou em seguida durante 1 minuto à velocidade 15.000rpm, e re-centrifugada durante 2 minutos na mesma velocidade. Por fim, foram adicionados 50uL de Água Milli Q à coluna, este volume mantido em contacto com a coluna durante 1 minuto à TA e centrifugado à velocidade de 15.000rpm.

3.2.5 Eletroforese em gel de agarose

A eletroforese em gel de agarose foi utilizada para verificar os produtos de reação do PCR e analisar o resultado das reações de restrição enzimático. A eletroforese em gel foi realizada com um gel de agarose de 0,8% (w/v), por um período de 1 hora com voltagem entre 90-100V. Após a eletroforese, o gel de agarose foi observado com recurso a uma lâmpada de UV e analisadas as bandas obtidas. O tamanho das bandas obtidas foi comparado com recurso ao marcador de DNA GeneRuler 1kb (Fermentas[®]).

3.2.6 Sequenciação

O processo de sequenciação foi realizado pela empresa GATC Biotech[®] no clone positivo à restrição enzimática diagnóstica. Como a sequência do gene CG11505 é

grande foi necessário utilizar *primers*, referenciados na tabela 7, que detetassem as regiões mais externas e a região interna da sequência. Os resultados da sequência foram analisados com recurso ao programa Genome Compiler® (www.genomecompiler.com).

| Primer | Sequência | Local de sequenciação |
|------------|--|-----------------------|
| Fw | TCAACTGCAACTACTGA | Externa |
| Rev | CATCAGTTCCATAGGTTG | Externa |
| Fw | ATGCGGCCGCATGTACCCATACGATGTTCCAGATTACGCTATGAACGGAGAT GCACTC | Interna |
| Rev | GGTCTAGATCAGAAGAGGCCGCGCCTTGTGGT | Interna |

Tabela 7 - Lista de primers utilizados na sequenciação do vetor de expressão *pUAST-CG11505*

3.2.7 Análise bioinformática

A análise do alinhamento das proteínas foi realizada com recurso à função *DIOPT-DRSC Integrative Ortholog Prediction* do site flyrnai.org. Os ortólogos da proteína LARP4 foram identificados através da colocação na secção *Enter Genes and/or Proteins* a designação LARP4 e selecionado na secção *Input Species: Homo sapiens (Human)* e na *Output Species: Drosophila melanogaster (Fly)* ou *Mus Musculus (Mouse)* e, posteriormente, analisados os alinhamentos das proteínas. Todas as análises dos alinhamentos das proteínas foram re-analisadas em Março de 2018.

3.2.8 Linhas de sobre-expressão e knock-down

As moscas transgénicas que expressão CG11505, sob controlo do sistema de expressão GAL/UAS, foram gerados através da integração *Random* do transgene (P {w [+]= UAS-Larp} no genoma de moscas com *background* de olhos brancos, na empresa BestGene Inc,. Posteriormente, a transmissão da linhagem germinativa foi analisada através da observação de moscas que não apresentassem olhos

brancos na geração F2. A análise da sua inserção no genoma foi determinada com recurso ao uso de cromossomas balanceados (CyO e RM3, Sb), pelo nosso colaborador Dr. Aaron Voigt, que identificou nas dez linhas independentes que o UAS-CG11505 estava inserido no cromossoma X, no 2º e 3º cromossomas.

Na análise de sub-expressão da proteína CG11505 foi adquirido na Bloomington Drosophila Stock Center, a linha BL#64937: *y[1] sc[*] v[1];P{y[+t7.7] v[+t1.8]=Trip.HMC05810}attP40*.

3.2.9 Análise funcional da proteína CG11505 em Drosophila

Em colaboração com o Dr. Aaron Voigt foram realizadas análises funcionais de sobre- e sub-expressão da CG11505 no modelo animal de *Drosophila melanogaster* para analisar o efeito da CG11505 nos neurónios do Sistema nervoso (*elav*-Gal4) e na retina (*GMR*-Gal4).

As dez linhas transgénicas foram cruzadas com moscas contendo o *driver* *elav*[GS]-Gal4 (*y[1] w[*];P{w[+mC]=elav-Switch.O}GSG301*). Ao atingir o estadio adulto, as moscas foram alimentadas com a hormona, por um período de 6 dias, e analisados os níveis de expressão da proteína CG11505, por western Blot, através da deteção do tag HA nas dez linhas transgénicas.

Para analisar o efeito no olho, as moscas transgénicas e a mosca RNAi_CG11505 foram cruzadas com o *driver* *GMR*-Gal4 (*w[*];P{w[+mC]=GAL4- ninaE.GMR}12*) e foram observadas as alterações fenotípicas macroscopicamente bem como ao microscópio para análise da estrutura e número dos omatídeos.

IV. Resultados

A SMA é uma doença neurodegenerativa resultante da diminuição da quantidade da proteína SMN devido à deleção ou mutação do gene SMN1. Não se sabe muito acerca da fisiopatologia da doença nem existe tratamento para a mesma. Porém, modificadores genéticos da expressão do Smn têm sido estudados como possíveis alvos terapêuticos. O laboratório de acolhimento identificou anteriormente a proteína LARP4, bem como outras proteínas da mesma família, como um potencial regulador da expressão do mRNA dos genes SMN1 e 2 em células humanas (Oliveira, 2010). Para além disso, o seu homólogo – CG11505 – foi identificado como um potencial modificador positivo da SMA, no modelo animal de *Drosophila melanogaster* (Sen et al., 2013).

Desta forma, o presente trabalho pretende investigar a função desta proteína na regulação da expressão do SMN em linhas celulares, bem como caracterizar a sua expressão e o seu papel no sistema nervoso central de organismos modelo. Especificamente, pretendeu-se caracterizar o impacto funcional do aumento ou diminuição da expressão da proteína LARP4 em linhas celulares e no sistema nervoso de *Drosophila melanogaster*, bem como caracterizar pela primeira vez a expressão desta proteína nos tecidos neuronais e musculares de mamíferos.

4.1 Análise bioinformática da homologia entre CG11505 e LARP4

A proteína CG11505 foi identificada, no modelo de SMA de *Drosophila melanogaster*, como sendo um modificador positivo do SMN (A. Sen et al., 2013). A *Drosophila melanogaster* apresenta duas proteínas LARP, designadas por *larp* (CG42551 ou CG14066 ou CG14065) e CG11505. Desta forma pretendeu-se realizar o alinhamento das sequências das proteínas CG11505 e LARP4, com o intuito de analisar se a proteína CG11505 é a mais semelhante com a proteína humana LARP4.

O alinhamento das sequências das proteínas CG11505 e LARP4 revelou que estas apresentam uma identidade de 27% (221/813) e uma semelhança de 43% (246/813) entre si. Relativamente aos domínios, verificou-se que o domínio *LARP_4_5_like* da proteína CG11505 tem grande semelhança com domínio *LARP_4* da proteína LARP4, que é de 62% (46/74). Por outro lado, o domínio *RRM*

4.2 Localização celular da proteína LARP4

A proteína LARP4 encontra-se localizada no citoplasma da célula, em células HeLa (Yang et al., 2011). Porém, não se sabe a co-localização celular das proteínas LARP4 e SMN. Desta forma recorreu-se às linhas celulares HEK293, SK-N-Be(2)c e GM03813 para analisar a localização celular da proteína LARP4 e a sua co-localização com a proteína SMN por imunofluorescência.

No caso da linha celular HEK293, que são células de rim de embrião humano, observou-se que a proteína LARP4 localiza-se no citoplasma e que a proteína SMN apresenta uma localização tanto ao nível do núcleo como do citoplasma da célula. O padrão de localização das duas proteínas é semelhante no citoplasma, apresentando uma distribuição difusa (Figura 11A).

A localização celular das mesmas proteínas nas células SK-N-Be(2)c, que é uma linha celular de neuroblastoma, revela ser a mesma, relativamente à linha celular HEK293, isto é, a proteína LARP4 encontra-se localizada no citoplasma e a SMN encontra-se localizada tanto no núcleo como no citoplasma (Figura 11B).

Por outro lado, na linha celular GM03813 (SMA fibroblastos), derivada de um doente com SMA tipo I (com duas cópias do gene SMN2), a localização espacial da proteína LARP4 é diferente da que foi observado nas restantes linhas celulares previamente referidas. Para além de uma distribuição citoplasmática, a proteína LARP4 apresenta também uma distribuição no núcleo (Figura 11C).

Desta forma pode-se concluir que a distribuição da localização da proteína LARP4 é predominantemente citoplasmática. Porém, quando há diminuição dos níveis de SMN, no caso da linha celular GM03813, o padrão de distribuição da proteína LARP4 difere-se das restantes linhas celulares estudadas.

Considerando que esta linha celular apresenta uma diminuição da quantidade da proteína SMN, pretendeu-se analisar se a localização da proteína LARP4 é similar numa linha celular de fibroblasto sem alterações nos níveis do SMN. Para tal, optou-se pela utilização da linha celular VA-13 (Figura 11D). Desta forma, foi possível verificar que a distribuição espacial da proteína LARP4 é similar à observada nas linhas celulares HEK293 e SK-N-Be(2)c. Portanto, este resultado

sugere que a distribuição nuclear de LARP4 é específica da linha celular GM03813, podendo estar associado aos/ocorrer em consequência dos baixos níveis de SMN. Relativamente à proteína SMN na linha celular GM03813, esta encontra-se localizada no núcleo e no citoplasma (Figura 11C). O padrão de distribuição da proteína SMN é similar ao que foi observado nas restantes linhas celulares (Figura 11A-B, setas), sendo o seu padrão típico de distribuição na célula. Nesta linha celular, observou-se que as proteínas LARP4 e SMN não se encontram co-localizadas no núcleo (Figura 11C, seta laranja).

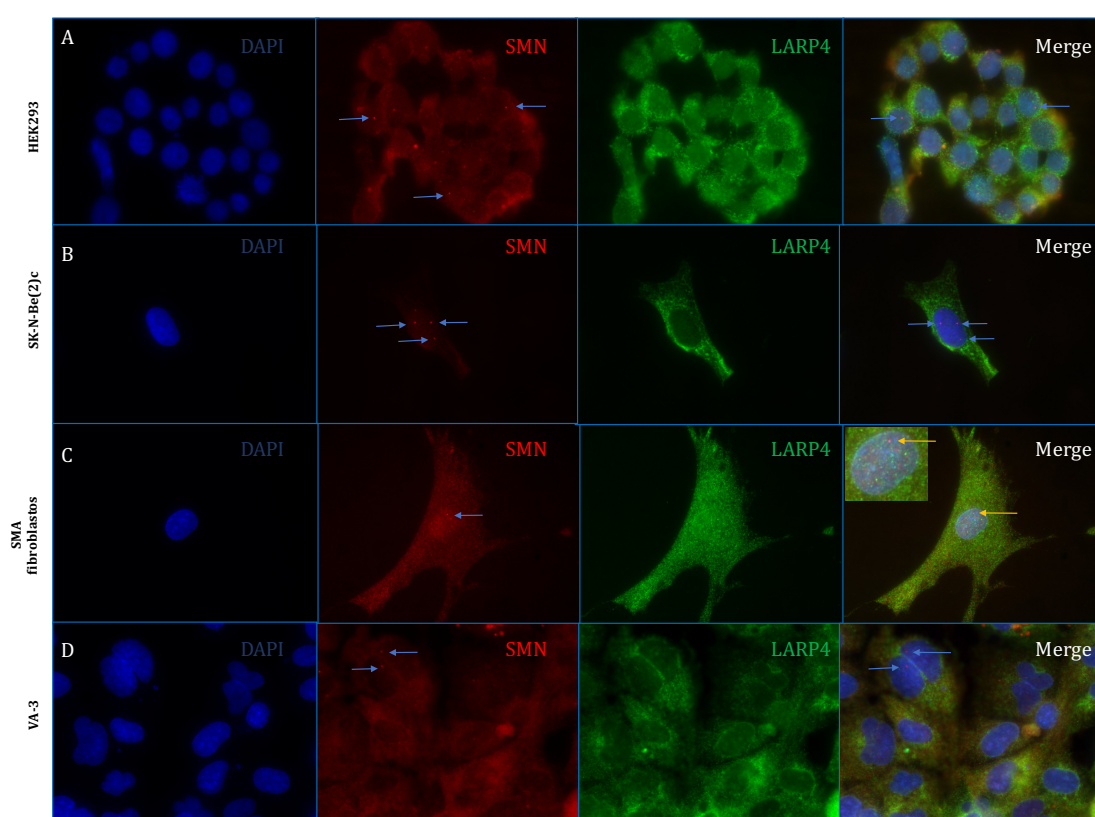


Figura 11 – Imunofluorescência das proteínas SMN, LARP4 e coloração de DNA (DAPI) nas linhas celulares HEK293 (A), SK-N-BE(2)c (B), GM03813 (C) e VA-13 (D). A proteína SMN tem uma distribuição nuclear (seta azul). Na linha celular GM03813 observa-se que as proteínas LARP4 e SMN não se co-localizam no núcleo (seta laranja).

O padrão de distribuição típico da proteína SMN, no núcleo, é em duas estruturas sub-nucleares: o *gems* e os Corpos de Cajal. A função do *gems* não é muito bem esclarecido e esta estrutura encontra-se próxima ou em estreita relação com os Corpos de Cajal. Por outro lado, uma das funções dos Corpos de Cajal está relacionada com o metabolismo do RNA, nomeadamente o splicing do pré-mRNA. Desta forma, pretendeu-se verificar se a proteína LARP4 tem uma localização nos Corpos de Cajal, através da marcação com Coilina, visto ser o marcador específico para esta estrutura sub-nuclear (Raška et al., 1991). Através da observação dos resultados, verificou-se que a LARP4 e a Coilin não se localizam no mesmo local no núcleo (Figura 12, seta amarela).

Por isso, podemos concluir que quando há diminuição dos níveis de SMN a distribuição da proteína LARP4 é alterada, verificando-se uma localização tanto no núcleo como no citoplasma. No núcleo, esta proteína não se encontra localizada nas duas estruturas sub-nucleares que contém a proteína SMN.

Portanto, estes resultados sugerem que a proteína LARP4 não terá uma interacção direta com o SMN. Para além disso, as duas proteínas não se encontram nas mesmas estruturas sub-nucleares - *Gems* e Corpos de Cajal. Contudo, verifica-se que há uma alteração da sua localização da LARP4 quando há alteração dos níveis de SMN.

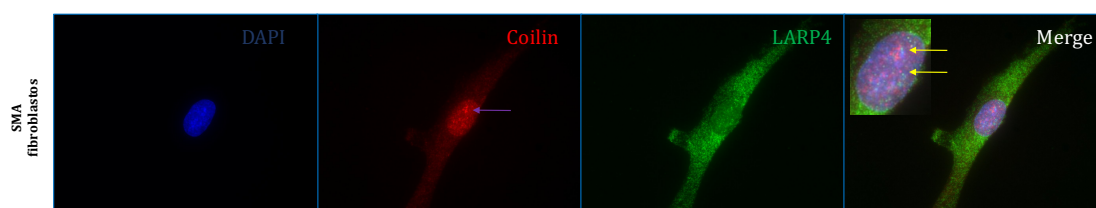


Figura 12 -Imunofluorescência das proteínas Coilin, LARP4 e coloração de DNA (DAPI) na linha celular GM03813. A proteína Coilin tem uma distribuição nuclear, nos Corpos de Cajal (seta roxa) e a LARP4 e a Coilin não se localizam no mesmo local (seta amarela).

4.3 Análise histológica da localização da proteína LARP4 nas junções neuromusculares em embriões de ratinho

Laviolette e seus colaboradores (2005) identificaram a proteína *larp* (CG42551 ou CG14066 ou CG14065) como estando relacionada com o crescimento das junções neuromusculares em *Drosophila melanogaster*. A proteína *larp* tem um homólogo humano designada por LARP1 e que também foi identificada por Oliveira (2010) como uma proteína que se liga à região 3'UTR do mRNA do SMN1/2.

Desta forma pretendeu-se analisar se a proteína LARP4 apresenta uma localização nas junções neuromusculares em embriões de ratinho (E13.5 e E15.5) antes e após a maturação destas estruturas, em colaboração com o grupo de Biologia do Desenvolvimento da Professora Solveig Thorsteinsdottir na Faculdade de Ciências da Universidade de Lisboa.

Primeiramente foi analisado o alinhamento da proteína *Larp4*, de ratinho, e a proteína humana LARP4 para analisar a sua homologia. A proteína homóloga de LARP4 em ratinho, a *Larp4*, é composta por 720 aminoácidos, apresentando uma identidade de 83% (603/727) e uma semelhança de 89% (647/727) quando comparando a LARP4. Relativamente aos domínios, verificou-se que o domínio *LARP_4* da proteína LARP4 tem a mesma (71/74, 96%) identidade do domínio *LARP_4* da proteína *Larp4*. O domínio *RRM_LARP4* da proteína LARP4 apresenta a mesma identidade (73/76, 96%) do domínio *RRM_LARP4* da proteína *Larp4* (Figura 13).

Portanto, pode-se concluir que as duas proteínas apresentam um elevado grau de conservação dos seus domínios e, desta forma, uma similaridade na sua função e que pode ser possível a utilização do anticorpo anti-LARP4 utilizado previamente na imunofluorescência nas linhas celulares humanas.

| | | | |
|-------|-----|--|-----|
| Human | 14 | QQVASKGTGLNPNKAVQEIAPGNTDAPVTHGTESSWHEIAATSGAHPEVSFNTGNAELSEDI | 78 |
| Mouse | 6 | EQVTSKGTGLNPNKAVQEIIPSGNPDGTPV---TEPSWHETAATSGSHPE-----GHTELSDEM | 61 |
| Human | 79 | CKEYEVMYSSSCETTRNTTGIEESTDGMILGPEDLSYQIYDVSGESNSAVSTEDLKECLKKQLEF | 143 |
| Mouse | 62 | CKEYEVMYSPSCETTRNTADVEESADGMILGPEDLSYQLYDVSGESSAISTEDLKECLKKQLEF | 126 |
| Human | 144 | CFSRENLSKDLYLISQMDSQDFIPIWTVANMEEIKKLTDPDILEVLRSSPMVQVDEKGEKVRP | 208 |
| Mouse | 127 | CFSRENLSKDLYLISQMDSQDFVPIWTVANMEEIKKLTTNTDLILEVLRSSPMVQVDEKGEKVRP | 191 |
| Human | 209 | SHKR CIVILREIPETTPIEEVKGLFKSENCPKVISCEFAHNSNWYITFQSDTDAQQAFKYLREEV | 273 |
| Mouse | 192 | SHKR CIVILREIPETTPVEEVKALFKNENCPKVISCEFAHNSNWYITFQSDTDAQQAFKYLREEV | 256 |
| Human | 274 | KTFQKGPIMARIKAINITFAKNGYRLMDSSIIYSHPIQTQAQYASPVFMQPVYNPHQQYSVYSIVP | 338 |
| Mouse | 257 | KTFQKGPIMARIKAINITFAKNGYRLMDSSMYTQPIQTPTQYPSPVFMQPVYNP-QQYSVYSLVP | 320 |
| Human | 339 | QSWSPNPTPYFETPLAPFPNGSFVNGFNSPGSYKTNAAMNMGRPFQKNRVKQFRSSGGSEHST | 403 |
| Mouse | 321 | QSWSPSPAPYFETPLAPFPNGSFVNGFSSPGSYKTNAAMNMGRPFQKNRVKPHFRSSGGSEHST | 385 |
| Human | 404 | EGSVSLGDGQLNRYSSRNFFAERHNPTVTGHQEQTLYLQKETSTLQVEQNGDYGRGRRTLFRGRRR | 468 |
| Mouse | 386 | EGSVSLGDGPLSRSSSRIFLSERHNPTVTGQQEQTYLPKEAPILQMEQNGDFGRGRRTLFRGRRR | 450 |
| Human | 469 | REDDRISRPHPSTAESKAPTPKFDLLASNFPPLPGSSSRMPGELVLENRMSDVVKGVYKEKDNEE | 533 |
| Mouse | 451 | RDDDRIPRPQPAATEAKAPTPKFDLLATNFPPLPGSSSRVPDELGLENRMSDVVKGVCREKDS | 515 |
| Human | 534 | LTISCPVPA-DEQTECTSAQQLNMSTSSPCAAELTALSTTQKEKDLIEDSSVQKDGLNQTTIPVS | 597 |
| Mouse | 516 | VRVSCPVAEDGQTDCTSA-PLSISPSPPCTAEPVVLSTTQEQDQMEDSAVPKDTLNPVAVPVS | 579 |
| Human | 598 | PPSTTKPSRASTASPCNNNINAATAVALQEPRKLSYAEVCQKPPKEPSSVLVQPLRELRSNVVSP | 662 |
| Mouse | 580 | SPTATKPSANTASPTCSNINPPRAVALQEPRKLSYAEVCQKPPKEPSPVLVQPLRELRSNAVSP | 644 |
| Human | 663 | TKNEDNGAPENSVEKPEKPEARASKDYSGFRGNIIPRGAAGKIREQRRQFSHRAIPQGVTRRNG | 727 |
| Mouse | 645 | TRNEENGAPKPEKPEKPEKTRASKDHSYSGFRGNTIPRGAAGKIREQRRQFSHRAIPQGVTRRNG | 709 |
| Human | 728 | KEQYVPPRSPK | 738 |
| Mouse | 710 | KEQYVPPRSPK | 720 |

Figura 13 - Alinhamento das proteínas LARP4 humana e Larp4 de ratinho, com recurso ao site *flyrnai.org*

A localização da proteína LARP4 nas junções neuromusculares foi analisada na região pre-sináptica com recurso aos anticorpos primários anti-Neurofilamento e anti-Sinaptofisina (Figura 14A e B) e na região pós-sináptica utilizando a toxina α -Bungarotoxina (Figura 14D e E). Porém, não foi possível a identificação da localização da proteína Larp4 em comparação com o controlo negativo, nos tecidos de embriões E13.5 e E15.5.

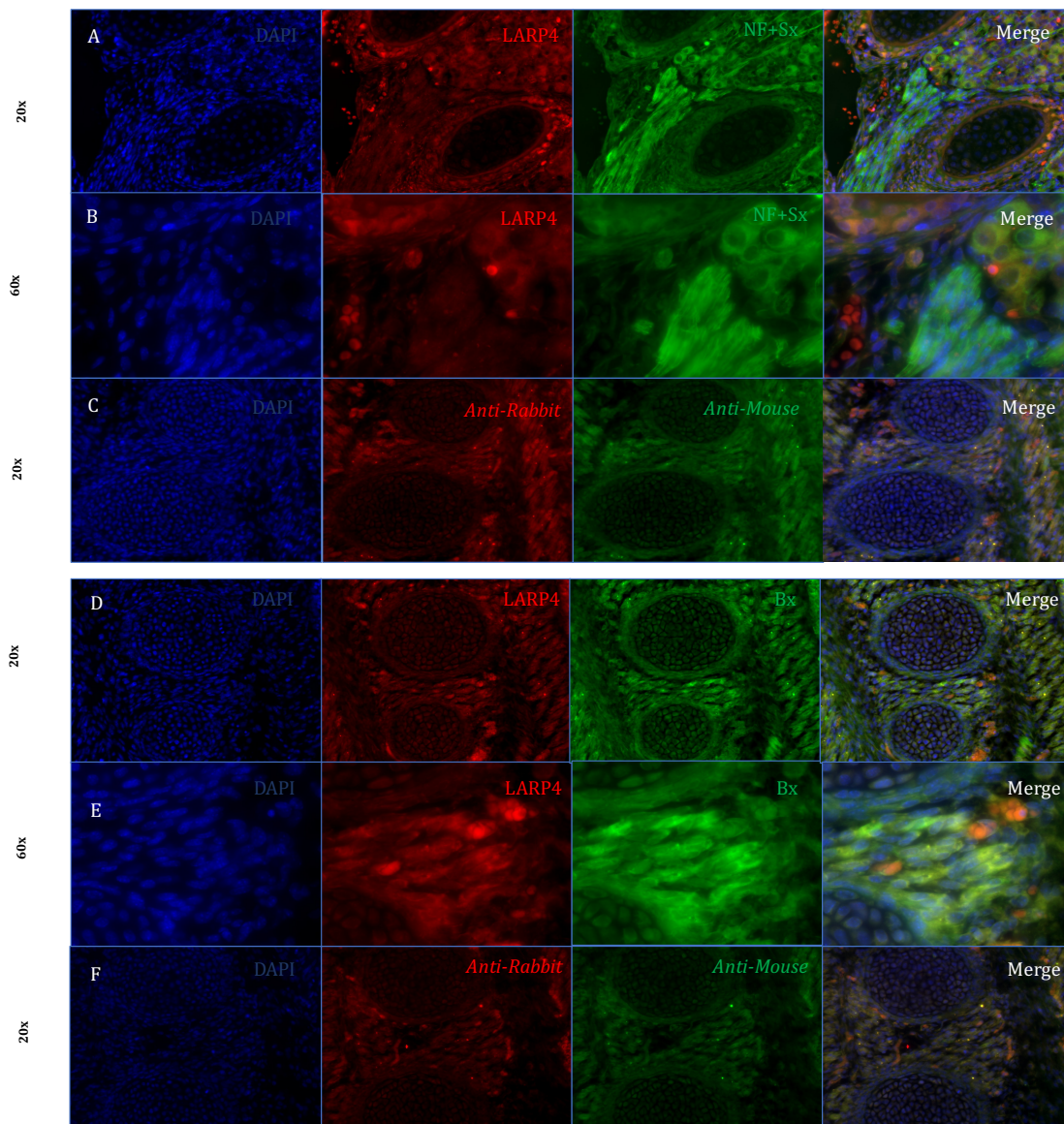


Figura 14 – Imunohistoquímica da localização da LARP4 na junção neuromuscular em embrião de ratinho, E15.5, na região pré-sináptica (A-B) com recurso aos anticorpos primários anti-Neurofilamento (NF) e anti-Sinaptofisina (SF) e região pós-sináptica (D-E) através da toxina α -Bungarotoxina (Bx) com ampliação de 20x e 60x. As imagens C e F representam o controlo negativa da técnica de imunohistoquímica para as regiões pré- e pós-sináptica, respetivamente.

Visto ter sido utilizado um anticorpo anti-LARP4 que não estava referenciado para utilização em ratinho, colocou-se a hipótese de o resultado anterior poder dever-

se ao facto de este anticorpo não estar a ser reconhecido. Desta forma, procurou-se demonstrar se o anticorpo identifica e localiza a proteína em linhas celulares de ratinho, recorrendo a um ensaio de imunocitoquímica.

A observação ao microscópio das linhas celulares 3T3 e C2C12, foi possível verificar que o anticorpo identifica a proteína Larp4 em células de ratinho e apresenta o mesmo padrão de distribuição observado previamente nas linhas celulares humanas (Figura 15).

Por isso, estes resultados podem sugerir que a proteína Larp4 não se encontra localizada nas regiões pré- e pós-sináptica, durante o desenvolvimento das junções neuromusculares em ratinho. Portanto, será necessária uma avaliação em outros momentos de desenvolvimento do ratinho para perceber se a proteína Larp4 está presente ou não nas junções neuromusculares.

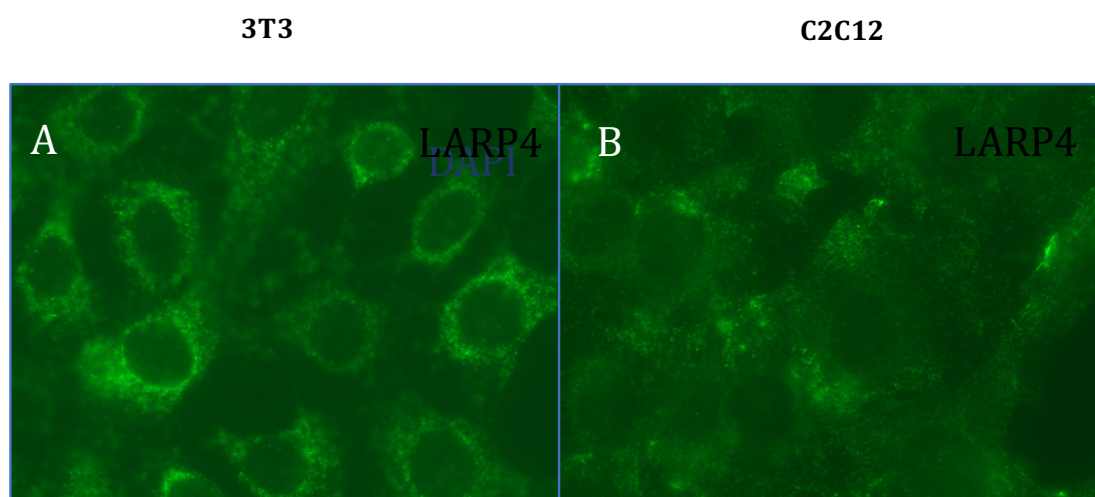


Figura 15 – Imunofluorescência da proteína LARP4 linhas celulares 3T3 (A) e C2C12 (B).

4.4 Análise do papel da LARP4 na regulação da expressão da proteína SMN em modelos celulares

As regiões não traduzidas do mRNA do SMN1/2 apresentam um papel fundamental na regulação pos-transcricional. A região 3'UTR do mRNA tem um efeito positivo na regulação do mRNA do SMN1/2; porém, a região 5'UTR do mRNA tem um efeito negativo na sua regulação (Gomes, 2010). Para além destas

observações, também se verificou que a região 3'UTR está associada à regulação da estabilidade do mRNA. Por isso, foram identificadas proteínas que se ligam à região 3'UTR do mRNA do SMN1/2 e verificou-se que uma das proteínas era a LARP4 (Oliveira, 2010). Sen e seus colaboradores (2013) identificaram a proteína CG11505, homóloga da LARP4 humana, como um modificador positivo do Smn. Desta forma, pretendeu-se analisar o efeito da expressão da LARP4 sob o gene repórter Luciferase-3'UTR do SMN1/2 e o Smn endógeno, em condições de sobre- e sub-expressão, ao nível da proteína e do RNA.

4.4.1 Efeito da expressão da LARP4 sob o gene repórter Luciferase-3'UTR do SMN1/2

O efeito da LARP4 sob o gene repórter Luciferase-3'UTR do SMN1 foi analisado através da avaliação dos níveis de luciferase e dos níveis de mRNA de luciferase, nas condições de sobre- e sub-expressão da LARP4.

Os resultados de sobre-expressão demonstraram que o aumento da expressão da LARP4 resulta no aumento significativo dos níveis de luciferase do gene repórter da região 3'UTR do mRNA do SMN1/2, na linha celular HEK293 ($p < 0,05$), validando os resultados obtidos previamente pelo colega do laboratório de acolhimento (Figura 16). Os níveis de mRNA da Luciferase, também foram analisados previamente, tendo demonstrado que não havia alteração nos níveis de expressão do mRNA.

Porém, o inverso não foi observado quando se induziu a diminuição dos níveis de LARP4. A linha celular HEK293 foi transfetada com RNA de interferência para LARP4, durante um período de 72 horas, e co-transfetada com os vetores de expressão e avaliados os níveis de expressão de LARP4 através de PCR em tempo-real (Figura 16B). Nas condições de sub-expressão da LARP4, os ensaios de luciferase demonstraram que, quando há diminuição dos níveis de LARP4, não há alteração nos níveis de luciferase do gene repórter Luciferase-3'UTR do mRNA do SMN1/2 ($p > 0,05$). Contudo, os níveis de mRNA de luciferase diminuem quando há diminuição dos níveis de LARP4 ($p < 0,05$), não tendo qualquer impacto nos níveis de luciferase.

Portanto, pode-se deduzir que a proteína LARP4 tem um efeito positivo sob o gene repórter Luciferase-3'UTR do mRNA do SMN1/2, sugerindo que a LARP4 tem um papel no aumento da tradução do gene repórter. Para além disso, a diminuição da LARP4 resulta na diminuição dos níveis de mRNA, possivelmente, por alteração da estabilidade do mRNA.

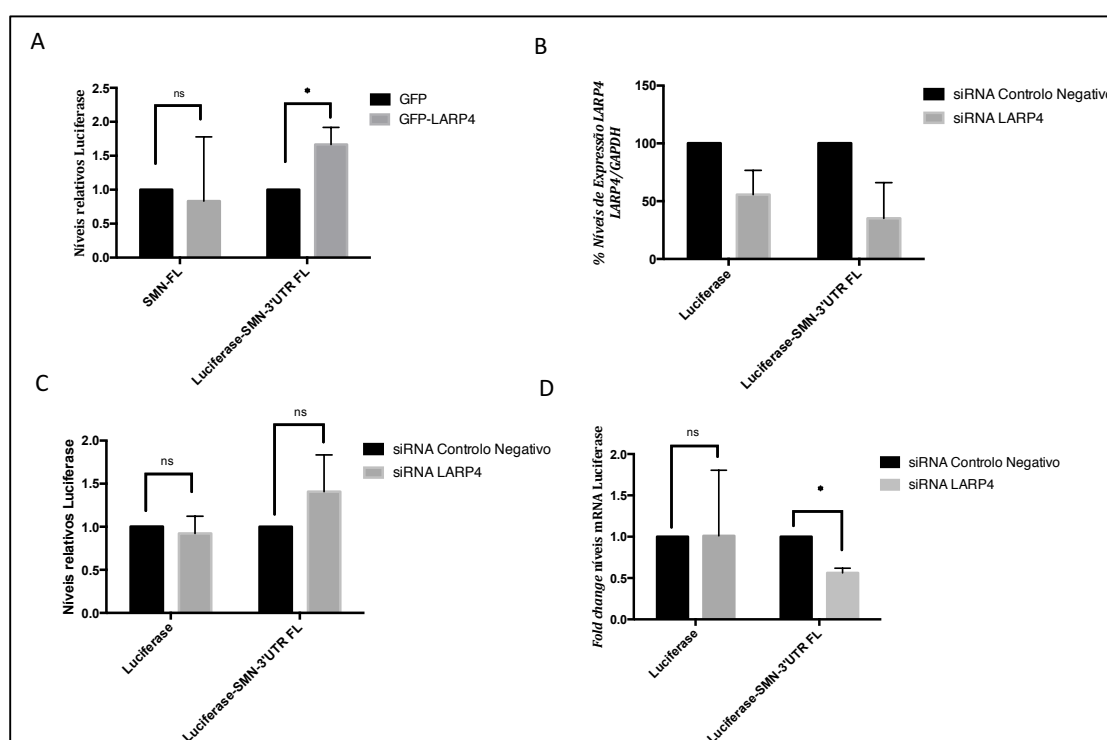


Figura 16 - Análise dos níveis de luciferase, na linha celular HEK293, após transfectar com os vetores de expressão GFP e GFP-LARP4 sob os genes repórter SMN-FL e Luciferase-SMN-3'UTR FL (A). (B) Gráficos que demonstram os níveis de expressão da LARP4 após interferência com siRNA LARP4. Análise dos níveis de luciferase (C) e dos níveis de mRNA de luciferase (D), na linha celular HEK293, após transfectar as células com Controlo Negativo siRNA e o siRNA LARP4 e avaliar o seu efeito nos genes repórter SMN-FL e Luciferase-SMN-3'UTR FL. Todos os dados estão expressos em média \pm desvio padrão de 3 amostras independentes. ns significa não significativa com $p > 0,05$ e * significa significância com $p < 0,05$, após aplicação do teste estatístico test t student.

4.4.2 Efeito da expressão da LARP4 sob o SMN endógeno

Na sequência da análise dos efeitos dos níveis da LARP4 na expressão do gene repórter Luciferase-3'UTR do SMN1/2 verificou-se que a LARP4 tem um efeito positivo na sua expressão e diminuindo os níveis de mRNA quando há diminuição dos níveis de LARP4. Desta forma, pretendeu-se analisar o efeito da LARP4 na expressão do Smn endógeno, nas condições de sobre- e sub-expressão em linhas celulares.

As linhas celulares HEK293 e SK-N-Be(2)c foram transfetadas com os vetores de expresso GFP e GFP-LARP4 e os níveis de expressão de LARP4 foram avaliados através de PCR em tempo real (Figura 17A e D, respetivamente.)

Na linha celular HEK293 não se verificou qualquer alteração nos níveis de expressão da proteína Smn endógena quando se aumentou os níveis de LARP4 em comparação com o controlo (GFP) ($p > 0,05$) (Figura 17B). Pelo contrário, na linha celular SK-N-Be(2)c observou-se um aumento dos níveis de expressão da proteína Smn ($p < 0,05$) (Figura 17E). Relativamente aos níveis de mRNA, não se observou variação nos níveis de mRNA em comparação com o controlo (GFP), nas duas linhas celulares ($p > 0,05$) (Figura 17C e F, respetivamente).

Desta forma pode-se concluir que o aumento dos níveis da proteína LARP4 tem um efeito positivo na expressão do Smn endógeno na linha celular SK-N-Be(2)c, sem interferir nos níveis de mRNA, como foi observado nos ensaios realizados com o gene repórter Luciferase-3'UTR do SMN1/2. Porém, este efeito não foi observado na linha celular HEK293 possivelmente porque os níveis da proteína LARP4 já se encontra nos seus níveis máximos podendo tornar-se tóxicos para a célula. Desta forma, estes dados revelam que o aumento da expressão da LARP4 resulta no aumento da tradução da proteína endógena Smn.

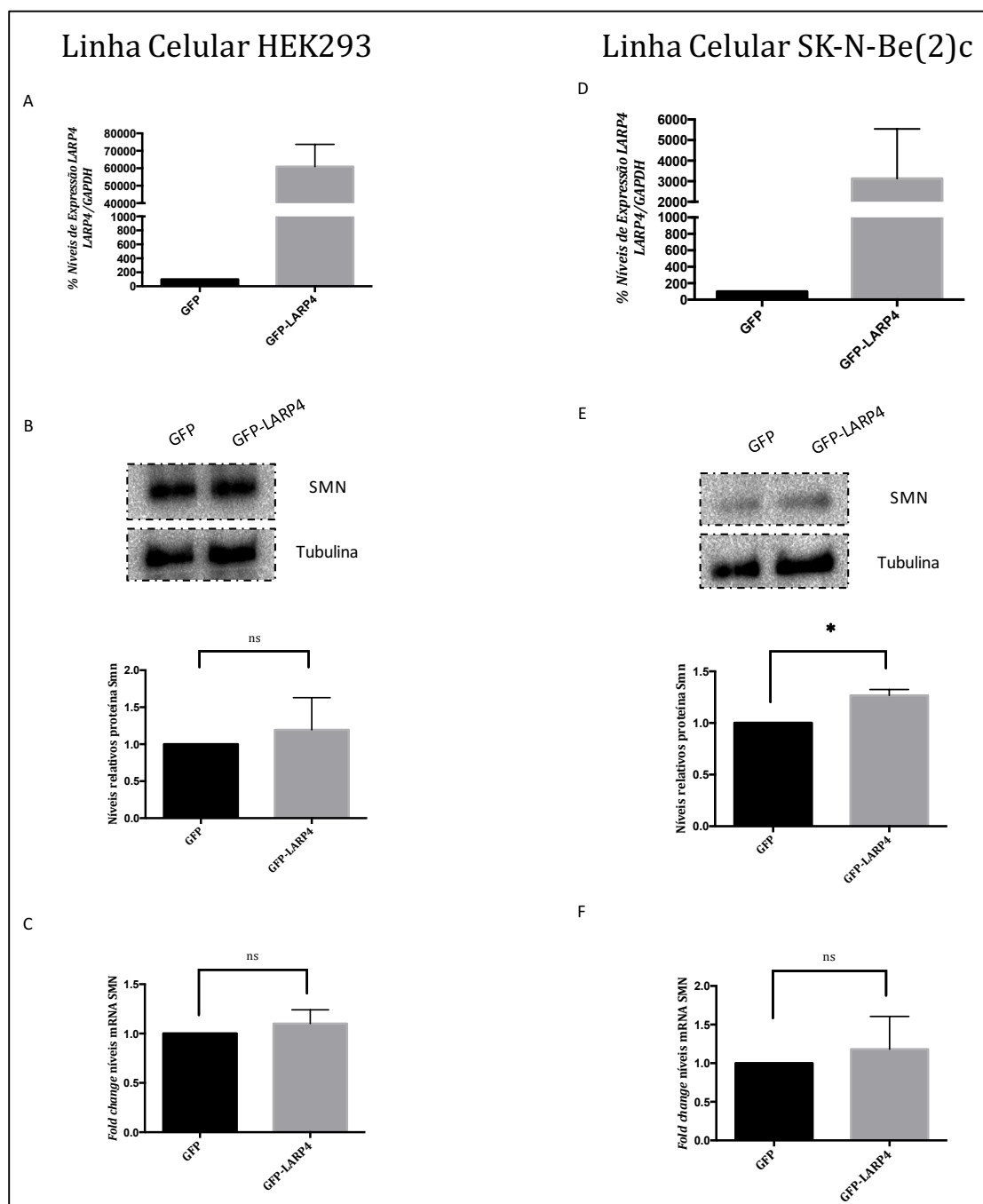


Figura 17 - Análise funcional da interação entre LARP4 e Smn endógena, após transferir as linhas celulares HEK293 e SK-N-BE(2)c com os vetores de expressão GFP e GFP-LARP4 e verificados os níveis de expressão através de PCR em tempo real (A) (B). Na linha celular HEK293 verificou-se que não há alteração nos níveis de expressão da proteína Smn endógena (B), nem dos níveis de expressão de mRNA do SMN endógeno (C). Na linha celular SK-N-BE(2)c observou-se um aumento de expressão da proteína Smn endógena (E), sem alteração nos níveis de expressão de mRNA do SMN endógeno (F). Todos os dados estão expressos em média±desvio

padrão de 3 amostras independentes. ns significa não significante e * significa significância com $p < 0.05$, após aplicação do teste estatístico *test t student*.

No que concerne aos ensaios de sub-expressão da LARP4, as linhas celulares HEK293 e SK-N-Be(2)c foram transfetadas com RNA de interferência para LARP4, durante 72 horas, e os níveis de expressão de LARP4 foram avaliados através de PCR em tempo-real, como observado nos gráficos representados na Figura 18.

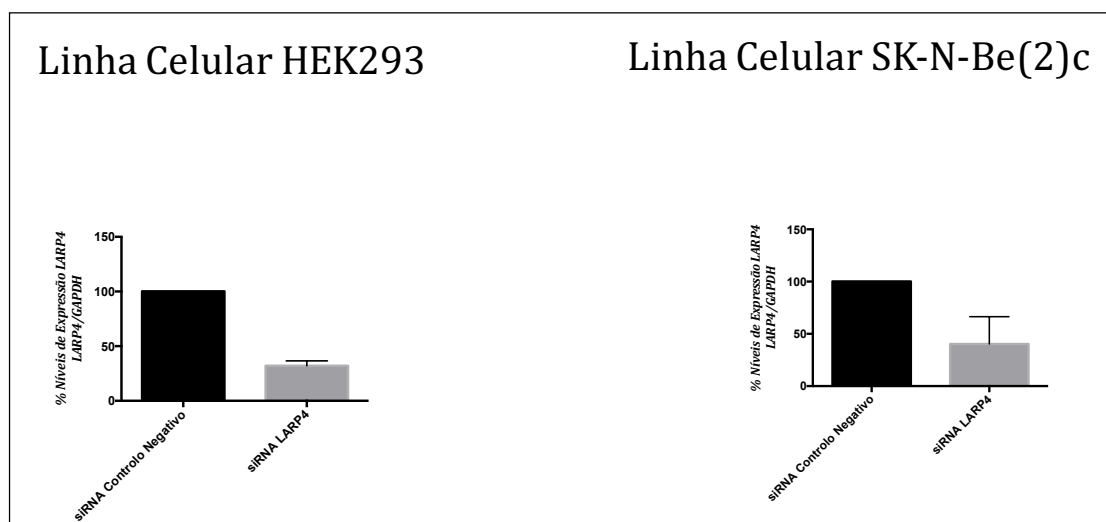


Figura 18 - Gráficos que demonstram os níveis de expressão da LARP4 após ensaios de interferência nas linhas celulares HEK293 e SK-N-Be(2)c

Após a confirmação da expressão da LARP4, observou-se que os níveis da proteína Smn endógena não se alteraram, nas duas linhas celulares ($p > 0,05$), como foi observado anteriormente nos ensaios de luciferase com o gene repórter Luciferase-3'UTR do SMN1/2 (Figura 20, em cima). Em contraste e como observado anteriormente, verificou-se uma diminuição dos níveis de mRNA do SMN endógeno, nas duas linhas celulares como verificado previamente nos ensaios com o gene repórter Luciferase-3'UTR do SMN1/2 ($p < 0,05$) (Figura 19, em baixo).

Portanto, os ensaios de sub-expressão revelam que a diminuição da LARP4 resulta na diminuição dos níveis do mRNA do SMN, possivelmente devido a uma alteração na estabilidade do mRNA.

Desta forma, pode-se concluir que a LARP4 é um modificador positivo na expressão do Smn como foi identificado anteriormente por Sen e seus colaboradores (2013), no modelo animal *Drosophila melanogaster*.

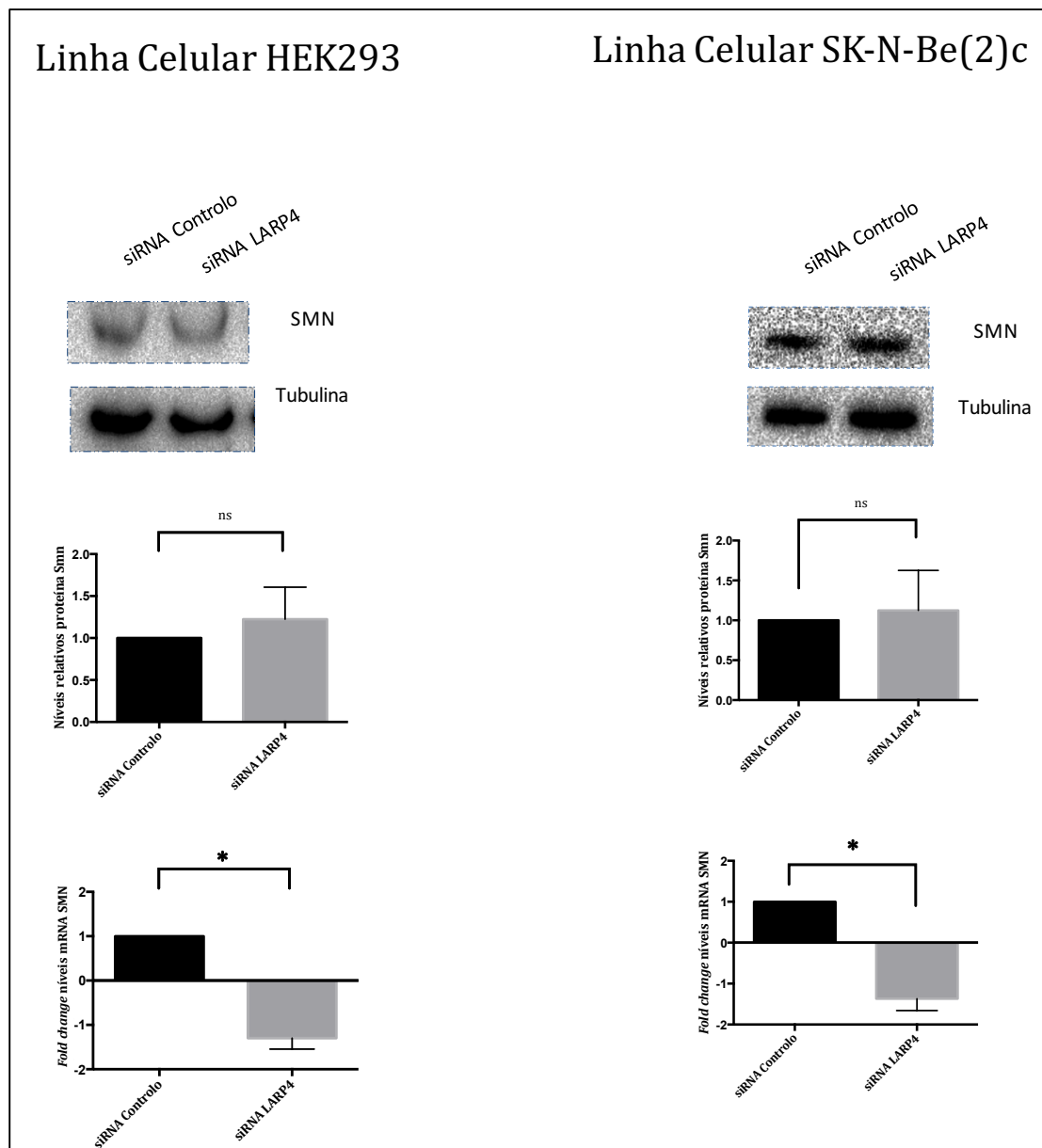


Figura 19 - Análise funcional da interação entre LARP4 e Smn endógeno, após transfetar as linhas celulares HEK293 e SK-N-Be(2)c com siRNA controlo negativo e siRNA LARP4, durante 72 horas. Na linha celular HEK293 verificou-se que não há alteração nos níveis de expressão da proteína Smn endógena (em cima, esquerda), mas uma diminuição dos níveis de expressão do mRNA do SMN endógeno (em baixo, esquerda). Na linha celular SK-N-Be(2)c não se observou alterações na expressão da proteína Smn endógena (em cima, direita), mas uma diminuição dos

*níveis de expressão de mRNA do SMN endógeno (em baixo, direita). Todos os dados estão expressos em média±desvio padrão de 3 amostras independentes. ns significa não significativa com $p > 0,05$ e * significa significância com $p < 0.05$, após aplicação do teste estatístico test t student.*

4.5 Estudo da função do homólogo de *Drosophila* da proteína LARP4 no sistema nervoso

Sen e seus colaboradores (2013) identificaram a proteína CG11505 como um modificador positivo da SMA, no modelo animal de *Drosophila melanogaster*. Para além disso, Laviolette e seus colegas (2005) identificou outra proteína da família das LARP's como modificadora das junções neuromusculares. Por isso, e visto não haver informação acerca do papel da proteína LARP4 foi imperativo analisar o efeito desta proteína no sistema nervoso, nas condições de sobre- e sub-expressão.

4.5.1 Clonagem do vetor de expressão pUAST CG11505

Para analisar a expressão do gene LARP4, em células de *Drosophila melanogaster*, foi necessário a obtenção de moscas transgênicas que sobre-expressam o gene de interesse, através do sistema UAS/GAL4. Este método consegue expressar um gene alvo num tecido específico, permitindo conhecer e analisar a função do gene alvo, neste caso conhecer a função e analisar o impacto no Sistema nervoso. Para isso, foi construído um vetor de expressão (pUAST_CG11505), tendo na sua constituição um tag de HA que permite identificar a correta inserção do vetor de expressão nas moscas transgênicas.

O primeiro passo na construção do vetor de expressão foi a obtenção do gene CG11505 com extremidades onde se encontrem locais de restrição para as enzimas de restrição de eleição (NotI e XbaI). Para isso, recorreu-se a uma reação em cadeia de polimerase, tendo como molde o vetor pDNR_Dual_CG11505 (AT27581). A análise do produto de PCR foi realizada através de uma eletroforese em gel de agarose para permitir analisar a integridade e o tamanho do produto e,

posteriormente, a sua excisão e purificação. O tamanho da banda do gene CG11505 esperado é de cerca de 2900 bp, como verificado na figura 20B.

O vetor pUAST e o produto de PCR sofreram um processo de restrição enzimática com as enzimas NotI e XbaI, para a obtenção de extremidade coesivas. Estas extremidades coesivas permite a ligação do gene ao vetor pUAST.

A reação de restrição enzimática foi analisada através de eletroforese em gel de agarose, permitindo verificar a eficiência da reação e validar o tamanho das bandas obtidas e se estavam de acordo com o esperado. O resultado obtido foi uma banda com cerca de 8904 bp que corresponde ao vetor pUAST linear e uma banda de cerca de 2900 bp que corresponde ao produto de PCR do gene CG11505. Estas duas bandas foram excisadas e o seu DNA purificado, para sofrerem o processo de ligação com o intuito de obter um constructo final pUAST_CG1150 e , posteriormente, transformadas em bactérias competentes (Figura 20C).

Os clones obtidos foram isolados e analisados com recurso a reação de restrição enzimática com a enzima de restrição EcoRI. O que se esperava era a obtenção de um conjunto de três bandas que correspondiam a um total de 11807bp. Portanto, de acordo com o esperado, o que se verificou nos clones positivos foi a presença de três bandas de DNA com os tamanhos de 10221bp, 1173bp e 413bp. O único clone positivo foi sequenciado, tendo-se confirmado a sequência codificante clonada (Anexo I).

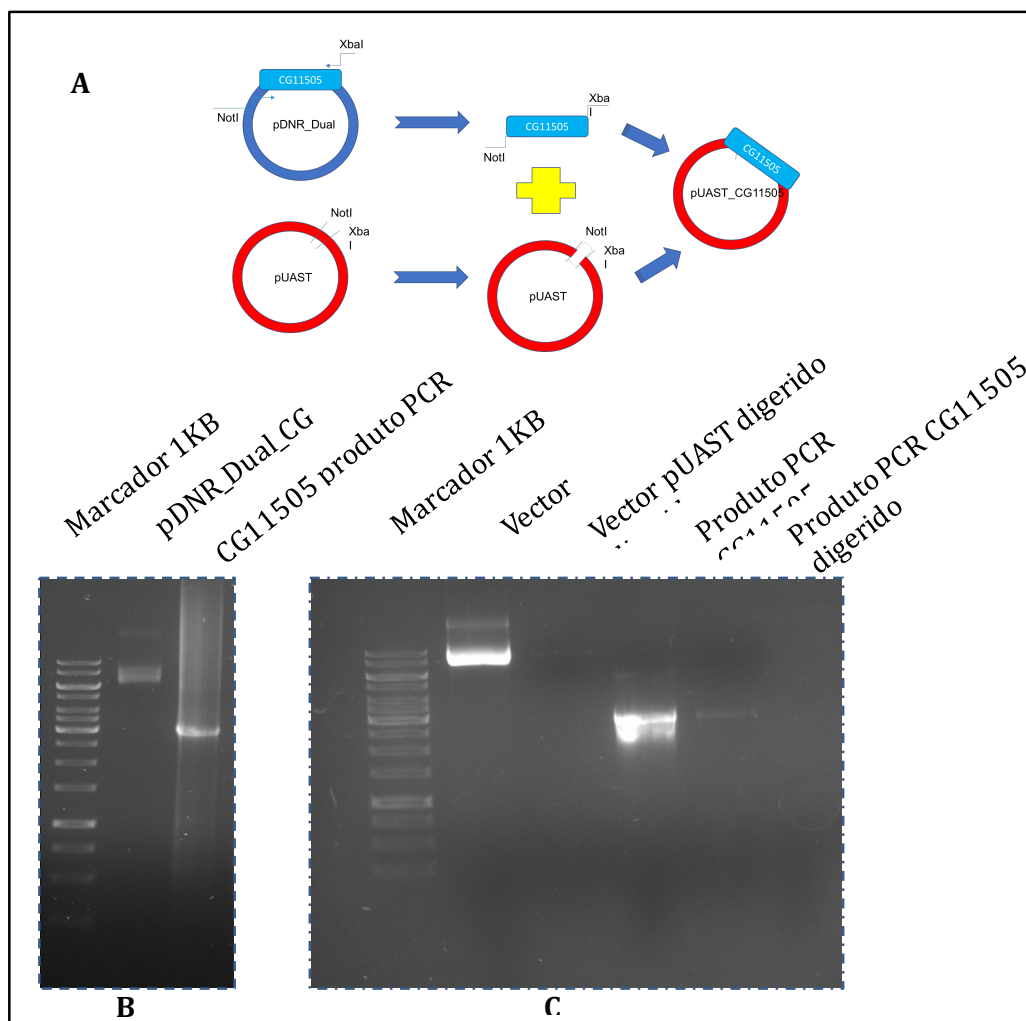


Figura 20 - Diagrama do método de clonagem do construto pUAST_CG11505 (A).
Imagens de eletroforese em geral da reacção de PCR do gene CG11505 (B) e da
reacção de restrição do vetor pUAST e do produto de PCR do gene CG11505, com as
enzimas de restrição NotI e XbaI (C).

4.5.2 Análise do impacto da sobre-expressão do gene CG11505 em *Drosophila melanogaster*

O vetor de expressão construído foi enviado para a empresa BestGene Inc para a produção de moscas transgênicas, em colaboração com o Doutor Aaron Voigt do Departamento de Neurologia da Universidade de RWTH Aachen (Alemanha), tendo sido obtidas dez linhas de moscas transgênicas balanceadas.

Posteriormente, as moscas transgênicas balanceadas foram cruzadas com diferentes genótipos contendo *drivers* para a expressão da proteína recombinante de forma específica nos neurónios do sistema nervoso (*elav-Gal4*) e na retina (*GMR-Gal4*) (Figura 22A). Os níveis de expressão da proteína CG11505 foram analisados através de Western-Blot, com recurso ao anticorpo anti-HA, visto o vetor de expressão conter o tag HA (Figura 21A).

Após a análise da expressão proteica verificou-se que as moscas transgênicas apresentavam diferentes níveis de expressão da proteína CG11505, sendo que a linha 2 apresenta a menor expressão e a 10 a maior expressão (Figura 21B). A amostra proveniente da linha transgênica número 4 foi degradada, não tendo sido possível avaliar o seu nível de expressão.

Verificou-se que, na maioria dos casos, a expressão pan-neuronal da CG11505 é letal, estando relacionada com as moscas transgênicas com maior expressão da proteína CG11505, exceto a mosca transgênica número 9. Pelo contrário, as moscas transgênicas com menor expressão da proteína CG11505 são viáveis (Figura 22A).

No que concerne às alterações fenotípicas nos olhos das moscas, também, são dependentes dos níveis de expressão da proteína CG11505. Porém, quando se observa macroscopicamente os olhos das moscas transgênicas verifica-se que há uma alteração da superfície externa do olho e alteração da pigmentação entre as moscas transgênicas com diferentes níveis de expressão (Figura 22B). Contudo, quando os olhos são observados ao microscópio, observou-se que o número de omatídeos reduzem ou desaparecem quanto maior é a expressão de CG11505 (dados não apresentados).

Desta forma, as moscas transgênicas que apresentam maior expressão de CG11505 apresentam maior alteração fenotípica ao nível do olho e na longevidade demonstrando que não têm um papel apenas nos neurónios da retina, mas nos restantes neurónios também.

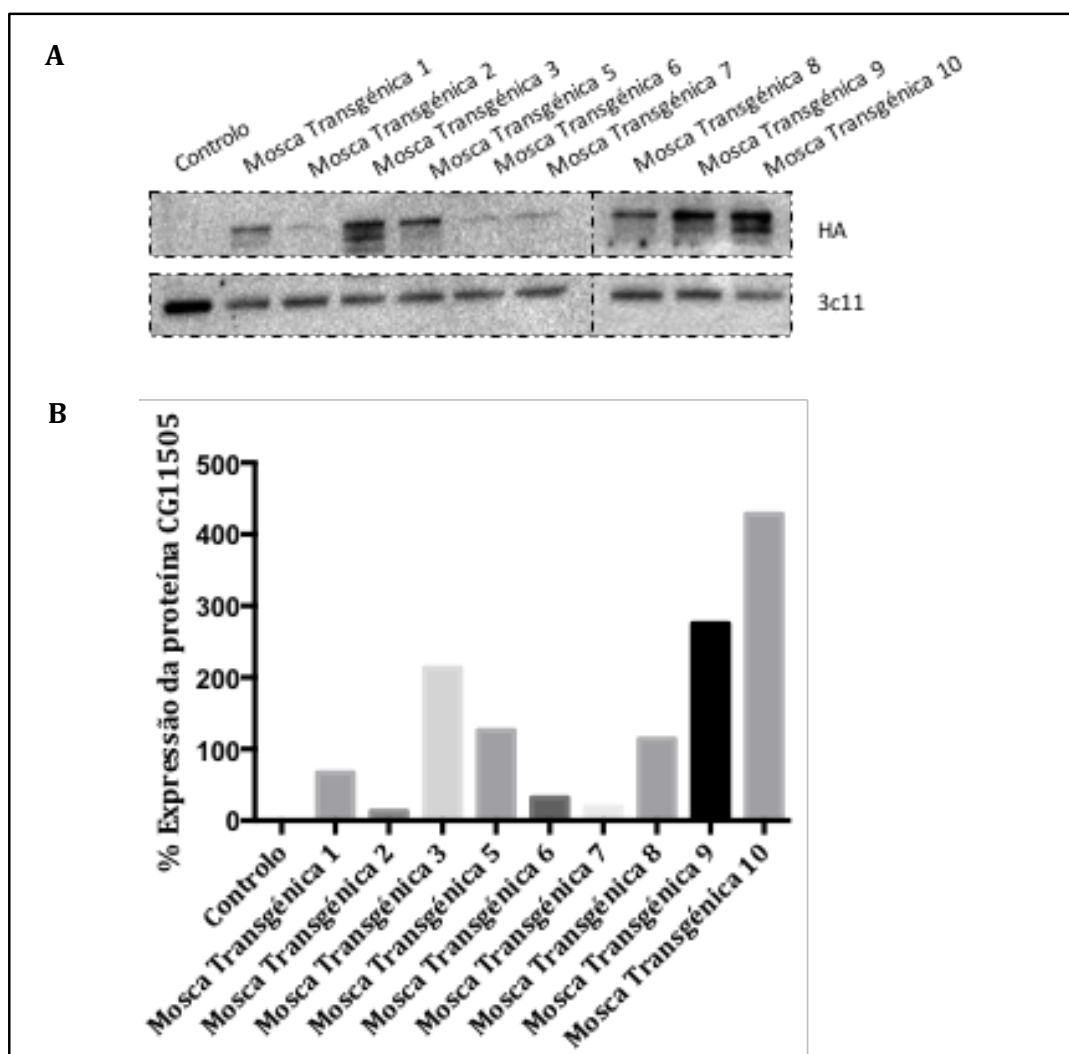


Figura 21 – Quantificação relativa da expressão da proteína CG11505 em moscas transgênicas. **A)** Western-Blot das amostras de proteína das moscas transgênicas balanceadas obtidas, identificando a expressão da proteína CG11505, que se encontra acoplado a um tag de HA, através do recurso do anticorpo anti-HA. **B)** Gráfico da quantificação da expressão da proteína CG11505 nas moscas transgênicas, tendo como controlo a proteína 3c11.

| | Mosca Transgênica | Cromossoma | Cruzamento | |
|----------|--|------------------|------------|-----------|
| | | | elav-Gal4 | GMR-Gal4 |
| A | #1 | Terceiro | Letal | REP |
| | #2 | Terceiro | Viável | nothing |
| | #3 | Terceiro | Letal | ✓ |
| | #4 | Terceiro | Letal | Letal (?) |
| | #5 | Segundo | Letal | REP |
| | #6 | Terceiro | ✓ | REP |
| | #7 | Terceiro | Viável | REP |
| | #8 | Terceiro | ✓ | ✓ |
| | #9 | Terceiro (letal) | Viável | ✓ |
| | #10 | Segundo (letal) | Semi-letal | REP |
| B | | | | |
| | Controlo Mosca transgênica 1 Mosca transgênica 1.2 Mosca transgênica 5 Mosca transgênica 6 | | | |
| | | | | |
| | Mosca transgênica 6.2 Mosca transgênica 7 Mosca transgênica 7.2 Mosca transgênica 10 | | | |

Figura 22 – Resultados dos cruzamentos das diferentes moscas transgênicas com os drivers *elav-GAL4* e *GMR-Gal4* (A) e imagens de olhos das moscas transgênicas que têm impacto na retina dos olhos (B)

4.5.3 Análise do impacto da sub-expressão do gene CG11505 em *Drosophila melanogaster*

Tendo-se verificado que o aumento da expressão da proteína CG11505 tem um impacto no sistema nervoso, foram adquiridas moscas transgênicas que expressam um RNA de interferência contra a proteína CG11505 no *Bloomington Drosophila Stock Center*, em colaboração com o Doutor Aaron Voigt do Departamento de Neurologia da Universidade de RWTH Aachen (Alemanha), com o intuito de perceber o efeito da diminuição da sua expressão no sistema nervoso, mais especificamente ao nível do olho. Para isso, as moscas transgênicas RNAi-CG11505 foram cruzadas com moscas que contém o *driver* para a expressão da proteína recombinante na retina (GMR-Gal4).

Verificou-se alteração do fenótipo do olho das moscas transgênicas com o RNA de interferência contra o CG11505, em comparação com o controlo, com uma alteração na pigmentação e do brilho, bem como o aparecimento de regiões pretas no olho que são indicadores de degenerescência (Figura 23). Por isso, estas observações sugerem que a diminuição da expressão de CG11505 resulta na degenerescência dos neurónios da retina e com um papel importante no sistema nervoso.

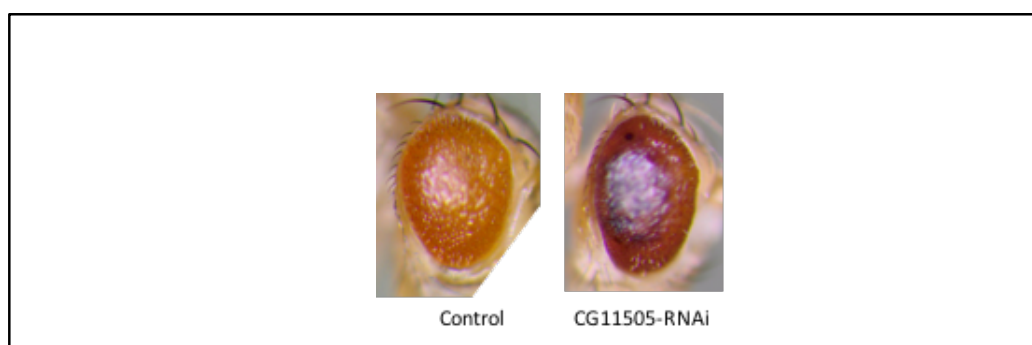


Figura 23 - Imagens do fenótipo nos olhos das moscas controlo e da mosca CG11505-RNAi

V. Discussão

A LARP4 foi identificada como uma *RNA binding protein* que interage com a região 3'UTR do mRNA do SMN1/2 (Oliveira, 2010) e como um modificador positivo do SMN (A. Sen et al., 2013). Desta forma, a LARP4 ao interagir com a região 3'UTR do mRNA do SMN1/2 poderá ter um efeito positivo na regulação da expressão do SMN. Portanto, o presente estudo pretendeu caracterizar o papel da LARP4 no sistema nervoso central e o seu efeito na expressão do SMN. Os resultados demonstraram que a LARP4 tem um efeito positivo na regulação da expressão do SMN.

A proteína LARP4 foi identificada no citoplasma nas linhas celulares HEK293 e SK-N-BE(2)c, como observado por Yang e seus colaboradores (2011). Por outro lado, na linha celular GM03813, fibroblastos derivados de um doente com SMA com duas cópias do gene SMN2, a distribuição da proteína é diferente da observada anteriormente. Nesta linha celular observou-se que a localização da proteína LARP4 encontra-se tanto no citoplasma como no núcleo das células. No núcleo, a sua localização não é co-localizada com a proteína SMN. A proteína SMN encontra-se localizada tanto no citoplasma como no núcleo, estando localizada nesta região da célula nas estruturas designadas por *Corpos de Cajal* e por *Gems*. Porém, através da identificação específica dos *Corpos de Cajal*, pelo anticorpo anti-Coilin, verificou-se que a proteína LARP4 também não se encontra nesta estrutura sub-nuclear.

Desta forma, os resultados observados sugerem que a diminuição da quantidade da proteína SMN na célula, devido a deleção ou mutação do gene SMN1 como no caso da patologia SMA, gera uma modificação do padrão de distribuição da proteína LARP4 na célula. Para além disso, sugerem também que, quando há diminuição dos níveis de SMN, a proteína LARP4 e SMN encontram-se no núcleo, mas não se encontram co-localizadas e que a LARP4 também não se encontra co-localizada com a proteína Coilina. Desta forma, quando há diminuição dos níveis de expressão de SMN, a proteína LARP4 encontra-se tanto no citoplasma como no núcleo mas não localizada nas estruturas sub-nucleares: *Gems* e *Corpos de Cajal*. A base de dados *Atlas Human Protein* referencia que a quantidade da proteína LARP4 e de SMN é diferente nos diversos tecidos do corpo humano e em diversas linhas celulares humanas. Esta informação parece apontar para que a proteína LARP4 apresente diferentes níveis de expressão nas diversas células e linhas

celulares, podendo apresentar diferentes limiares de tolerância nas diversas células do corpo humano. A proteína SMN também apresenta este mesmo padrão, em que diferentes células do corpo humano apresentam diferentes níveis da sua expressão. Porém, poderá haver uma ligação relativamente à expressão das duas proteínas, visto que nas células e linhas celulares em que os níveis de expressão da proteína LARP4 são mais elevados, também são os níveis da proteína SMN. O inverso é igualmente verificado.

A análise funcional foi realizada nas linhas celulares HEK293 e SK-N-Be(2)c, que apresentam diferentes quantidades de proteína LARP4 e SMN, tendo sido aumentados e diminuídos os níveis de expressão da LARP4 e analisado o seu impacto no SMN endógeno. A linha celular HEK293 é uma linha celular embrionária de rim que não apresenta característica de células neuronais, ao contrário da linha celular SK-N-Be(2)c, que é uma linha celular neuronal.

De acordo com os resultados obtidos verificou-se que o aumento dos níveis de expressão da proteína LARP4 tem um efeito positivo nos níveis de expressão da proteína Smn endógena. Desta forma, sugerem-se que a LARP4 poderá ter um papel na regulação da tradução da proteína SMN, resultado que está de acordo com as investigações de Yang e seus colaboradores (2011). Os autores mostram que o aumento da expressão da LARP4 está relacionado com o aumento da síntese proteica e, desta forma, com a regulação da tradução através da estabilidade do mRNA na linha celular HeLa. Porém, este resultado só foi observado na linha celular SK-N-Be(2)c, e não na linha celular HEK293. Contudo, em ambas as linhas celulares, não houve alterações nos níveis de expressão de mRNA do SMN.

No entanto, os dados obtidos nos ensaios de luciferase, realizados na linha celular HEK293, são contraditórios em relação ao que foi observado na análise da proteína Smn endógena na mesma linha celular. Os ensaios de luciferase têm por base a análise dos níveis de luciferase, na linha celular HEK293, após transfeção das células com um vetor que continha a região 3'UTR do mRNA do SMN1/2 e GFP-LARP4. Estes resultados demonstraram que quando as células HEK293 são transfetadas com o vetor que contém a região 3'UTR do SMN-Luciferase, juntamente com o aumento da expressão da LARP4, os níveis de luciferase aumentam indicando um aumento da tradução da luciferase. Para além disso, consegue-se perceber que o aumento dos níveis de luciferase deve-se à interação

da LARP4 com a região 3'UTR do mRNA do SMN1/2 (Oliveira, 2010), e que esta região tem um efeito positivo na regulação dos níveis de luciferase.

Porém, a LARP4 foi identificada como RBP na região 3'UTR do mRNA do SMN1/2 nos extratos da linha celular SK-N-BE(2)c, e não na linha celular HEK293 (Oliveira, 2010). Esta observação sugere, tal como referenciado na base de dados *Atlas Human Protein*, que a LARP4 apresenta vários níveis de expressão nas diferentes células e linhas celulares, demonstrando que os locais de ligação das regiões 3'UTR do mRNA do SMN1/2 podem estar ocupados de forma permanente com esta RBP, não tendo qualquer impacto no SMN se houver aumento da expressão da LARP4. Portanto, pode sugerir-se que a quantidade de proteína LARP4 é superior na linha celular HEK293 e que a região a que se liga o LARP4 na região 3'UTR do mRNA do SMN1/2 se encontra completamente saturada, e que o aumento da LARP4 não tem qualquer benefício para a célula, podendo mesmo o seu aumento tornar-se tóxico. Esta hipótese poderá ser apoiada com os resultados obtidos pelos ensaios de luciferase na linha celular HEK293, em que se aumentaram os níveis de expressão da região 3'UTR do mRNA do SMN1/2, bem como os níveis de LARP4, tendo-se observado o mesmo resultado obtido na linha celular SK-N-BE(2)c. Para além disso, estes dados também podem sugerir que a LARP4 apresenta um comportamento diferente em células neuronais e não-neuronais. Portanto, pode-se concluir que a LARP4 se liga à região 3'UTR do mRNA do SMN1/2, e tem um papel na regulação na tradução do SMN e possivelmente sem um efeito no mRNA do SMN1/2 visto não haver alterações nos níveis de expressão do mRNA de luciferase.

Por outro lado, quando se induz a diminuição dos níveis de expressão de LARP4 verifica-se uma diminuição dos níveis de mRNA do SMN1/2, sem qualquer tipo de efeito na tradução da proteína Smn endógena, através de observação por Western-Blot. Desta forma, os resultados observados sugerem que, em qualquer das linhas celulares estudadas, a diminuição dos níveis de LARP4 provoca uma diminuição dos níveis de expressão do mRNA do SMN1/2 possivelmente devido à instabilidade do mRNA e, por conseguinte, a sua degradação; sem ter um efeito negativo na tradução da proteína Smn endógena.

Estudos anteriores demonstraram que a proteína LARP4 é uma RBP que se liga à cauda poli(A) no citoplasma assim como à proteína PABPC através da interação

com uma variação do motivo PABP – PAM2 - que influencia a tradução e a estabilidade de alguns transcritos de mRNA (Yang et al., 2011). Para além disso, estudos prévios no laboratório de acolhimento demonstraram que a proteína LARP4 interage com a região 3'UTR do mRNA do SMN1/2 (Oliveira, 2010). A ligação da proteína PABPC resulta na estabilidade da cauda poli(A), impedindo que esta seja degradada por exonucleases no sentido 3'→5'. Desta forma, sugere-se que poderá haver uma cooperação entre as proteínas LARP4 e PABPC na região 3'UTR do mRNA do SMN1/2 e na cauda poli(A), permitindo indução da tradução do SMN pela interacção da proteína PABC com os fatores de iniciação eIF4G e eIF4E. Para além disso, também se sabe que a proteína LARP4 interage com o recetor ativador da proteína quinase C1 (RACK1) (Yang et al., 2011), sendo esta proteína responsável pela interacção com várias moléculas de sinalização e pelo recrutamento da proteína quinase C que ativa o ribossoma (Nilsson, Sengupta, Frank, & Nissen, 2004). Dessa forma, a ativação da RACK1 induz a tradução através da fosforilação do fator de iniciação 6, bem como de proteínas que se encontram associadas ao mRNA (Nilsson et al., 2004).

Portanto, a indução da tradução da proteína Smn endógena através do aumento da expressão da LARP4 poderá ser devido à interacção desta proteína com a PABPC e com a RACK1. Por outro lado, a diminuição da expressão LARP4 poderá induzir instabilidade do mRNA do SMN1/2, sem qualquer efeito na tradução da proteína endógena Smn mantendo os níveis essenciais para a célula.

No modelo animal de *Drosophila melanogaster*, a alteração dos níveis de expressão da CG11505 demonstrou um impacto no sistema nervoso. Nomeadamente, na sobre-expressão da proteína verificou-se que a maioria das moscas transgénicas morriam, bem como demonstravam uma alteração fenotípica do olho bem como alteração na longevidade. Desta forma, o aumento da expressão da proteína CG11505 pode sugerir que seja tóxica para o organismo da *Drosophila melanogaster* e, consequentemente, para o sistema nervoso.

Porém, não só o aumento da expressão da proteína CG11505, como a sua sub-expressão no modelo animal de *Drosophila melanogaster* tem um efeito negativo no sistema nervoso. A diminuição da expressão da proteína CG11505 resultou na degenerescência dos neurónios da retina sugerindo a indução de neurodegeneração.

A proteína CG11505 foi identificada como importante no desenvolvimento das dendrites, em *Drosophila melanogaster* (Olesnicky et al., 2014), e sua diminuição é responsável pela diminuição dos ramos das dendrites dos neurónios motores, principalmente dos segundos e terceiros ramos, e dos ramos terminais em *C. elegans* (Antonacci et al., 2015). Desta forma, a proteína LARP4 parece ter um papel no desenvolvimento do sistema nervoso central, principalmente no desenvolvimento das dendrites, visto demonstrar que a diminuição do homólogo da LARP4 resultar na degenerescência dos neurónios da retina, em *Drosophila melanogaster*.

Desta forma pode-se concluir que a proteína LARP4 demonstra que tem um papel relevante no controlo da expressão do SMN1/2 e que a diminuição da sua expressão poderá ter um efeito neurodegenerativo nos neurónios.

VI. Conclusão e Perspetivas Futuras

O presente trabalho expandiu o novo conhecimento acerca do papel da LARP4 na regulação da expressão do SMN e o seu impacto no sistema nervoso. Desta forma, os dados demonstraram que a proteína LARP4 tem efeito positivo na expressão do Smn, aumentando dessa forma a sua expressão endógena quando os seus níveis de expressão aumentam. Por outro lado, quando há diminuição da expressão da LARP4, os níveis de mRNA do SMN1/2 diminuem sem alteração nos níveis da proteína endógena. Para além disso, no modelo animal de *Drosophila melanogaster*, a sobre e sub-expressão do CG11505 resulta em alterações fenotípicas nos olhos. Nomeadamente, quando há diminuição dos níveis de CG11505 verificou-se um processo de degenerescência nos neurónios da retina. Contudo, muito trabalho é necessário realizar para entender de que forma a proteína LARP4 regula os níveis de SMN. Desta forma, seria importante avaliar o efeito da expressão da LARP4 sobre a proteína PABPC e RACK1 para tentar esclarecer o modo de regulação do SMN através da alteração da expressão da LARP4. Para além disso, também seria pertinente analisar o efeito da sobre- e sub-expressão da LARP4 em modelos animais de SMA, nomeadamente de *Drosophila melanogaster*, com o intuito de perceber se no sistema *in vivo* a modelação da expressão da LARP4 apresenta um efeito positivo na expressão do SMN e na melhoria da longevidade e no fenótipo da doença, assim como, o seu impacto nos neurónios motores.

VII. Referências Bibliográficas

- Adams, M., Celniker, S., Holt, R., Evans, C., Gocayne, J., Amanatides, P., ... An, H.-J. (2000). The Genome Sequence of *Drosophila melanogaster*. *Science*, 287(2185), 2185–2195. <http://doi.org/10.1126/science.287.5461.2185>
- Alberts, B., Bray, D., Hopki, K., Johnson, A., Lewis, J., Raff, M., ... Walter, P. (2014). *Essential Cell Biology* (4^a). New York: Garland Science.
- Antonacci, S., Forand, D., Wolf, M., Tyus, C., Barney, J., Kellogg, L., ... Killian, D. J. (2015). Conserved RNA-Binding Proteins Required for Dendrite Morphogenesis in *Caenorhabditis elegans* Sensory Neurons. *G3: Genes/Genomes/Genetics*, 5(4), 639–653. <http://doi.org/10.1534/g3.115.017327>
- Arraiano, C., & Fialho, A. (2007). *O Mundo do RNA - Novos desafios e perspectivas futuras* (1^a Edição). Lisboa: LIDEL - Edições Técnicas, Lda.
- Azevedo, F. A. C., Carvalho, L. R. B., Grinberg, L. T., Farfel, J. M., Ferretti, R. E. L., Leite, R. E. P., ... Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology*, 513(5), 532–541. <http://doi.org/10.1002/cne.21974>
- Baioni, M. T. C., & Ambiel, C. R. (2010). Spinal muscular atrophy: diagnosis, treatment and future prospects. *Jornal de Pediatria*, 86(4), 261–270. <http://doi.org/10.2223/JPED.1988>
- Battaglia, G., Princivalle, A., Forti, F., Lizier, C., & Zeviani, M. (1997). Expression of the SMN gene, the spinal muscular atrophy determining gene, in the mammalian central nervous system. *Human Molecular Genetics*, 6(11), 1961–1971. <http://doi.org/10.1093/hmg/6.11.1961>
- Bayfield, M., Yang, R., & Maraia, R. (2011). Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta*, 1799(5–6), 365–378. <http://doi.org/10.1038/jid.2014.371>
- Bear, M., Connors, B., & Paradiso, M. (2016). *Neuroscience: Exploring the Brain*.
- Béchéde, C., Rostaing, P., Cisterni, C., Kalisch, R., La Bella, V., Pettmann, B., & Triller, A. (1999). Subcellular distribution of survival motor neuron (SMN) protein: Possible involvement in nucleocytoplasmic and dendritic transport. *European Journal of Neuroscience*, 11(1), 293–304. <http://doi.org/10.1046/j.1460-9568.1999.00428.x>
- Bentley, D. L. (2014). Coupling mRNA processing with transcription in time and space. *Nature Reviews Genetics*, 15(3), 163–175. <http://doi.org/10.1038/nrg3662.Coupling>
- Boon, K. L., Xiao, S., McWhorter, M. L., Donn, T., Wolf-Saxon, E., Bohnsack, M. T., ... Beattie, C. E. (2009). Zebrafish survival motor neuron mutants exhibit presynaptic neuromuscular junction defects. *Human Molecular Genetics*, 18(19), 3615–3625. <http://doi.org/10.1093/hmg/ddp310>
- Bowerman, M., Becker, C. G., Yáñez-Muñoz, R. J., Ning, K., Wood, M. J. A., Gillingwater, T. H., & Talbot, K. (2017). Therapeutic strategies for spinal muscular atrophy: SMN and beyond. *Disease Models & Mechanisms*, 10(8), 943–954. <http://doi.org/10.1242/dmm.030148>
- Bozzo, F., Mirra, A., & Carri, M. T. (2017). Oxidative stress and mitochondrial damage in the pathogenesis of ALS: New perspectives. *Neuroscience Letters*, 636, 3–8. <http://doi.org/10.1016/j.neulet.2016.04.065>
- Brand, A. H., & Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*

- (Cambridge, England), 118(2), 401–15. <http://doi.org/10.1101/lm.1331809>
- Briese, M., Esmaeili, B., Fraboulet, S., Burt, E. C., Christodoulou, S., Towers, P. R., ... Sattelle, D. B. (2009). Deletion of smn-1, the *Caenorhabditis elegans* ortholog of the spinal muscular atrophy gene, results in locomotor dysfunction and reduced lifespan, 18(1), 97–104. <http://doi.org/10.1093/hmg/ddn320>
- Burghes, A., & Beattie, C. (2009). Spinal Muscular Atrophy: Why do low levels of SMN make motor neurons sick? *Nature Review Neuroscience*, 10(8), 597–609. <http://doi.org/10.1038/nrn2670.Spinal>
- Burglen, L., Lefebvre, S., Clermont, O., Burlet, P., Viollet, L., Cruaud, C., ... Melki, J. (1996). Structure and Organization of the Human Survival Motor Neurone (SMN) Gene. *Gene*, 482(2), 292–297. <http://doi.org/10.1006/geno.1996.0147>
- Burvill, E. (2009). The current understanding of motor neuron disease. *Current Anaesthesia & Critical Care*, 20(5–6), 236–239. <http://doi.org/10.1016/j.cacc.2009.07.009>
- Butchbach, M. E. R. (2016). Copy Number Variations in the Survival Motor Neuron Genes: Implications for Spinal Muscular Atrophy and Other Neurodegenerative Diseases. *Frontiers in Molecular Biosciences*, 3(7), 1–10. <http://doi.org/10.3389/fmolb.2016.00007>
- Campanari, M.-L., García-Ayllón, M.-S., Ciura, S., Sáez-Valero, J., & Kabashi, E. (2016). Neuromuscular Junction Impairment in Amyotrophic Lateral Sclerosis: Reassessing the Role of Acetylcholinesterase. *Frontiers in Molecular Neuroscience*, 9(160), 1–8. <http://doi.org/10.3389/fnmol.2016.00160>
- Cartegni, L., Hastings, M. L., Calarco, J. A., de Stanchina, E., & Krainer, A. R. (2006). Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *American Journal of Human Genetics*, 78(1), 63–77. <http://doi.org/10.1086/498853>
- Casso, D., Ramírez-Weber, F. A., & Kornberg, T. B. (1999). GFP-tagged balancer chromosomes for *Drosophila melanogaster*. *Mechanisms of Development*, 88(2), 229–232. [http://doi.org/10.1016/S0925-4773\(99\)00174-4](http://doi.org/10.1016/S0925-4773(99)00174-4)
- Chan, Y. B. (2003). Neuromuscular defects in a *Drosophila* survival motor neuron gene mutant. *Human Molecular Genetics*, 12(>12), 1367–1376. <http://doi.org/10.1093/hmg/ddg157>
- Chang, H. C. H., Dimlich, D. N., Yokokura, T., Mukherjee, A., Kankel, M. W., Sen, A., ... Artavanis-Tsakonas, S. (2008). Modeling spinal muscular atrophy in *Drosophila*. *PLoS ONE*, 3(9), 1–18. <http://doi.org/10.1371/journal.pone.0003209>
- Charroux, B., Pellizzoni, L., Parkinson, R. A., Yong, J., Shevchenko, A., Mann, M., & Dreyfuss, G. (2000). Gemin4: A novel component of the SMN complex that is found in both gems and nucleoli. *Journal of Cell Biology*, 148(6), 1177–1186. <http://doi.org/10.1083/jcb.148.6.1177>
- Cooper, G., & Hausman, R. (2013). *The Cell: A Molecular Approach* (6^a). Sunderland.
- Covert, D. D., Le, T. T., McAndrew, P. E., Strasswimmer, J., Crawford, T. O., Mendell, J. R., ... Burghes, A. H. M. (1997). The survival motor neuron protein in spinal muscular atrophy. *Human Molecular Genetics*, 6(8), 1205–1214. <http://doi.org/10.1093/hmg/6.8.1205>
- Council, N. R. (2000). *Scientific Frontiers in Developmental Toxicology and Risk Assessment*. Washington, DC: The National Academies Press.
- Crick, F. H. (1970). Central Dogma of Molecular Biology. *Nature*, 227(6), 561–563.

- <http://doi.org/10.1002/9781118906545>
- Cronin, S., Hardiman, O., & Traynor, B. (2007). Ethnic variation in the incidence of ALS: a systematic review. *Neurology*, 68(13), 1002–1007. <http://doi.org/10.1212/01.wnl.0000285431.01005.67>
- D'Amico, A., Mercuri, E., Tiziano, F. D., & Bertini, E. (2011). Spinal muscular atrophy. *Orphanet Journal of Rare Diseases*, 6(1), 71. <http://doi.org/10.1186/1750-1172-6-71>
- Dharmadasa, T., Henderson, R. D., Talman, P. S., Macdonell, R. Al, Mathers, S., Schultz, D. W., ... M.C., K. (2017). Motor neurone disease: progress and challenges. *The Medical Journal of Australia*, 206(8), 357–362. <http://doi.org/10.5694/mja16.01063>
- Edens, B. M., Ajroud-Driss, S., Ma, L., & Ma, Y. C. (2015). Molecular mechanisms and animal models of spinal muscular atrophy. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1852(4), 685–692. <http://doi.org/10.1016/j.bbadis.2014.07.024>
- Eidem, T., Kugel, J., & Goodrich, J. (2016). Noncoding RNAs: Regulators of the Mammalian Transcription Machinery. *Journal of Molecular Biology*, 428(12), 2652–2659. <http://doi.org/10.1016/j.cogdev.2010.08.003>. Personal
- Elsheikh, B. H., & Kissel, J. T. (2014). *Spinal muscular atrophies. Neuromuscular Disorders in Clinical Practice* (1st ed., Vol. 9781461465). Elsevier B.V. http://doi.org/10.1007/978-1-4614-6567-6_21
- Feldkötter, M., Schwarzer, V., Wirth, R., Wienker, T. F., & Wirth, B. (2002). Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *American Journal of Human Genetics*, 70(2), 358–68. <http://doi.org/10.1086/338627>
- Ferraiuolo, L., Kirby, J., Grierson, A. J., Sendtner, M., & Shaw, P. J. (2011). Molecular pathways of motor neuron injury in amyotrophic lateral sclerosis. *Nature Reviews Neurology*, 7(11), 616–630. <http://doi.org/10.1038/nrneurol.2011.152>
- Filippi, M., Agosta, F., Abrahams, S., Fazekas, F., Grosskreutz, J., Kalra, S., ... Masdeu, J. C. (2010). EFNS guidelines on the use of neuroimaging in the management of motor neuron diseases. *European Journal of Neurology*, 17(4), 526–533. <http://doi.org/10.1111/j.1468-1331.2010.02951.x>
- Fischer, J., Giniger, E., Maniatis, T., & Ptashne, M. (1988). GAL4 activates transcription in *Drosophila*. *Nature*, 332(6167), 853–856. <http://doi.org/10.1038/332853a0>
- Fredericks, A. M., Cygan, K. J., Brown, B. a, Fairbrother, W. G., & Biology, C. (2015). RNA-Binding Proteins: Splicing Factors and Disease, 893–909. <http://doi.org/10.3390/biom5020893>
- Frugier, T., Tiziano, F. D., Cifuentes-Diaz, C., Miniou, P., Roblot, N., Dierich, a, ... Melki, J. (2000). Nuclear targeting defect of SMN lacking the C-terminus in a mouse model of spinal muscular atrophy. *Human Molecular Genetics*, 9(5), 849–858. <http://doi.org/10.1093/hmg/9.5.849>
- Gabanella, F., Butchbach, M. E. R., Saieva, L., Carissimi, C., Burghes, A. H. M., & Pellizzoni, L. (2007). Ribonucleoprotein assembly defects correlate with spinal muscular atrophy severity and preferentially affect a subset of spliceosomal snRNPs. *PLoS ONE*, 2(9), 1–12. <http://doi.org/10.1371/journal.pone.0000921>

- Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS Letters*, 582(14), 1977–1986. <http://doi.org/10.1016/j.febslet.2008.03.004>
- Gomes, A. L. (2010). *Characterization of post-transcriptional control mechanisms regulating SMN2 gene expression*. Faculdade de Ciências da Universidade de Lisboa.
- Griffiths, A., Wessler, S., Lewontin, R., & Carroll, S. (2008). *Introduction to Genetic Analysis* (9^a). New York: W.H. Freeman & Company.
- Gubitz, A. K., Feng, W., & Dreyfuss, G. (2004). The SMN complex. *Experimental Cell Research*, 296(1), 51–56. <http://doi.org/10.1016/j.yexcr.2004.03.022>
- Hales, K. G., Korey, C. A., Larracuenta, A. M., & Roberts, D. M. (2015). Genetics on the fly: A primer on the drosophila model system. *Genetics*, 201(3), 815–842. <http://doi.org/10.1534/genetics.115.183392>
- Han, D. D., Stein, D., & Stevens, L. M. (2000). Investigating the function of follicular subpopulations during Drosophila oogenesis through hormone-dependent enhancer-targeted cell ablation. *Development (Cambridge, England)*, 127(3), 573–83. <http://doi.org/10.1101/gad.6.9.1752>
- Harada, Y., Sutomo, R., Sadewa, A. H., Akutsu, T., Takeshima, Y., Wada, H., ... Nishio, H. (2002). Correlation between SMN2 copy number and clinical phenotype of spinal muscular atrophy: Three SMN2 copies fail to rescue some patients from the disease severity. *Journal of Neurology*, 249(9), 1211–1219. <http://doi.org/10.1007/s00415-002-0811-4>
- Hebert, M., Szymczyk, P., Shpargel, K., & Gregory Matera, A. (2001). Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein. *Genes and Development*, 15(20), 2720–2729. <http://doi.org/10.1101/gad.908401>
- Heigwer, F., Port, F., & Boutros, M. (2018). Rna interference (RNAi) screening in Drosophila. *Genetics*, 208(3), 853–874. <http://doi.org/10.1534/genetics.117.300077>
- Hentze, M. W., Castello, A., Schwarzl, T., & Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nature Reviews Molecular Cell Biology*, 19(5), 327–341. <http://doi.org/10.1038/nrm.2017.130>
- Hsieh-Li, H. M., Chang, J. G., Jong, Y. J., Wu, M. H., Wang, N. M., Tsai, C. H., & Li, H. (2000). A mouse model for spinal muscular atrophy. *Nat Genet*, 24(january), 66–70. <http://doi.org/10.1038/71709>
- Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., & Mohr, S. E. (2011). An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics*, 12. <http://doi.org/10.1186/1471-2105-12-357>
- Huang, K., & Luo, Y. (1983). Adult spinal muscular athrophy. *Journal of the Neurological Sciences*, 61(2), 249–259. [http://doi.org/10.1016/0022-510X\(83\)90009-6](http://doi.org/10.1016/0022-510X(83)90009-6)
- Hussain, R. H., Zawawi, M., & Bayfield, M. A. (2013). Conservation of RNA chaperone activity of the human La-related proteins 4, 6 and 7. *Nucleic Acids Research*, 41(18), 8715–8725. <http://doi.org/10.1093/nar/gkt649>
- Iascone, D., Henderson, C., & Lee, J. (2015). Spinal muscular atrophy: from tissue specificity to therapeutic strategies. *F1000Prime Reports*, 7(January), 9. <http://doi.org/10.12703/p7-09>
- Iyer, J., Wang, Q., Le, T., Pizzo, L., Grönke, S., Ambegaokar, S. S., ... Girirajan, S.

- (2016). Quantitative Assessment of Eye Phenotypes for Functional Genetic Studies Using *Drosophila melanogaster*. *G3: Genes, Genome, Genetics*, 6(5), 1427–1437. <http://doi.org/10.1534/g3.116.027060>
- Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., & Hudspeth, A. J. (2013). *Principles of Neural Science* (5^a). McGraw-Hill Education.
- Kashima, T., & Manley, J. (2003). A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nature Genetics*, 34(4), 460–463.
- Katahira, J. (2015). Nuclear export of messenger RNA. *Genes*, 6(2), 163–184. <http://doi.org/10.3390/genes6020163>
- Kazemi-Esfarjani, P., & Benzer, S. (2000). Genetic suppression of polyglutamine toxicity in *Drosophila*. *Science*, 287(5459), 1837–1840. <http://doi.org/10.1126/science.287.5459.1837>
- Kennerdell, J. R., & Carthew, R. W. (2000). Heritable gene silencing in *Drosophila* using double-stranded RNA. *Nature Biotechnology*, 18(8), 896–898. <http://doi.org/10.1038/78531>
- Kolb, S. J., Battle, D. J., & Dreyfuss, G. (2007). Molecular Function of the SMN complex, 990–994. Retrieved from <http://jcn.sagepub.com/content/22/8/990.full.pdf>
- Kolb, S., & Kissel, J. (2011). Spinal muscular atrophy: A timely review. *Archives of Neurology*, 68(8), 979–984. <http://doi.org/10.1001/archneurol.2011.74>
- Kolb, S., Sutton, S., & Schoenberg, D. (2010). RNA processing defects associated with diseases of the motor neuron. *Muscle Nerve*, 41(1), 5–17. <http://doi.org/10.1002/mus.21428.RNA>
- Kolovos, P., Knoch, T. A., Grosveld, F. G., Cook, P. R., & Papantonis, A. (2012). Enhancers and silencers: An integrated and simple model for their function. *Epigenetics and Chromatin*, 5(1), 1. <http://doi.org/10.1186/1756-8935-5-1>
- Kostova, F., Williams, V., Heemskerk, J., Iannaccone, S., DiDonato, C., Swoboda, K., & Maria, B. (2007). Spinal muscular atrophy: Classification, diagnosis, management, pathogenesis, and future research directions. *Journal of Child Neurology*, 22(8), 926–945. Retrieved from <http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L47308306%5Cnhttp://dx.doi.org/10.1177/0883073807305662>
- Kwak, H., & Lis, J. (2013). Control of Transcriptional Elongation. *Annual Review of Genetics*, 47, 483–508. <http://doi.org/10.1146/annurev-genet-110711-155440>.
- Lai, S. L., & Lee, T. (2006). Genetic mosaic with dual binary transcriptional systems in *Drosophila*. *Nature Neuroscience*, 9(5), 703–709. <http://doi.org/10.1038/nn1681>
- Lam, J. K. W., Chow, M. Y. T., Zhang, Y., & Leung, S. W. S. (2015). siRNA versus miRNA as therapeutics for gene silencing. *Molecular Therapy - Nucleic Acids*, 4(9), 1–20. <http://doi.org/10.1038/mtna.2015.23>
- Lambrechts, D., Robberecht, W., & Carmeliet, P. (2007). Heterogeneity in motoneuron disease. *Trends in Neurosciences*, 30(10), 536–544. <http://doi.org/10.1016/j.tins.2007.07.002>
- Laviolette, M. J., Nunes, P., Peyre, J. B., Aigaki, T., & Stewart, B. A. (2005). A genetic screen for suppressors of *drosophila* NSF2 neuromuscular junction overgrowth. *Genetics*, 170(2), 779–792. <http://doi.org/10.1534/genetics.104.035691>
- Le, T. T., Pham, L. T., Butchbach, M. E. R., Zhang, H. L., Monani, U. R., Coover, D. D.,

- ... Burghes, A. H. M. (2005). SMN Δ 7, the major product of the centromeric survival motor neuron (SMN2) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN. *Human Molecular Genetics*, 14(6), 845–857. <http://doi.org/10.1093/hmg/ddi078>
- Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., ... Melki, J. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*, 80(1), 155–165. [http://doi.org/10.1016/0092-8674\(95\)90460-3](http://doi.org/10.1016/0092-8674(95)90460-3)
- Li, D. K., Tisdale, S., Lotti, F., & Pellizzoni, L. (2014a). SMN control of RNP assembly: From post-transcriptional gene regulation to motor neuron disease. *Seminars in Cell and Developmental Biology*, 32, 22–29. <http://doi.org/10.1016/j.semcdb.2014.04.026>
- Li, D. K., Tisdale, S., Lotti, F., & Pellizzoni, L. (2014b). SMN control of RNP assembly: From post-transcriptional gene regulation to motor neuron disease. *Seminars in Cell and Developmental Biology*, 32, 22–29. <http://doi.org/10.1016/j.semcdb.2014.04.026>
- Liang, X., Mahato, S., Hemmerich, C., & Zelhof, A. (2016). Two temporal functions of Glass: ommatidium patterning and photoreceptor differentiation. *Developmental Biology*, 414(1), 4–20. <http://doi.org/10.1016/j.ydbio.2016.04.012>
- Liu, Q., & Dreyfuss, G. (1996). A novel nuclear structure containing the survival of motor neurons protein. *The EMBO Journal*, 15(14), 3555–65. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=451956&tool=pmcentrez&rendertype=abstract>
- Lloyd, T. E., & Taylor, J. P. (2010). Flightless Flies: *Drosophila* models of neuromuscular disease. *Annals of the New York Academy of Sciences*, 1184, E1–E20. <http://doi.org/10.1111/j.1749-6632.2010.05432.x>
- Logroscino, G., Traynor, B. J., Hardiman, O., Chió, A., Couratier, P., Mitchell, J. D., ... Beghi, E. (2008). Descriptive epidemiology of amyotrophic lateral sclerosis: New evidence and unsolved issues. *Journal of Neurology, Neurosurgery and Psychiatry*, 79(1), 6–11. <http://doi.org/10.1136/jnnp.2006.104828>
- Lorson, C., & Androphy, E. (2000). An exonic enhancer is required for inclusion of an essential exon in the SMA-determining gene SMN. *Human Molecular Genetics*, 9(2), 259–265. <http://doi.org/10.1093/hmg/9.2.259>
- Lorson, C., Hahnen, E., Androphy, E., & Wirth, B. (1999). A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proceedings of the National Academy of Sciences of the United States of America*, 96(11), 6307–11. <http://doi.org/10.1073/pnas.96.11.6307>
- Lue, N. F., Chasman, D. I., Buchman, a R., & Kornberg, R. D. (1987). Interaction of GAL4 and GAL80 gene regulatory proteins in vitro. *Molecular and Cellular Biology*, 7(10), 3446–3451. <http://doi.org/10.1128/MCB.7.10.3446>. Updated
- Lunn, M. R., & Wang, C. H. (2008). Spinal muscular atrophy. *Lancet*, 371(9630), 2120–33. [http://doi.org/10.1016/S0140-6736\(08\)60921-6](http://doi.org/10.1016/S0140-6736(08)60921-6)
- Mailman, M. D., Heinz, J. W., Papp, A. C., Snyder, P. J., Sedra, M. S., Wirth, B., ... Prior, T. W. (2002). Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 4(1), 20–6. <http://doi.org/10.1097/00125817-200201000-00004>
- Manfredi, G., & Xu, Z. (2005). Mitochondrial dysfunction and its role in motor

- neuron degeneration in ALS. *Mitochondrion*, 5(2), 77–87. <http://doi.org/10.1016/j.mito.2005.01.002>
- McGuire, S. E., Le, P. T., Osborn, A. J., Matsumoto, K., & Davis, R. L. (2003). Spatiotemporal Rescue of Memory Dysfunction in *Drosophila*. *Science*, 302(5651), 1765–1768. <http://doi.org/10.1126/science.1089035>
- McKean, S., Ross, J., Dressler, D., & Scheurer, D. (2017). *Principles and Practice of Hospital Medicine* (1^a). McGraw-Hill Education.
- McWhorter, M. L., Monani, U. R., Burghes, A. H. M., & Beattie, C. E. (2003). Knockdown of the survival motor neuron (Smn) protein in zebrafish causes defects in motor axon outgrowth and pathfinding. *Journal of Cell Biology*, 162(5), 919–931. <http://doi.org/10.1083/jcb.200303168>
- Merret, R., Martino, L., Bousquet-Antonelli, C., Fneich, S., Descombin, J., Billey, E., ... Deragon, J.-M. (2013). The association of a La module with the PABP-interacting motif PAM2 is a recurrent evolutionary process that led to the neofunctionalization of La-related proteins. *RNA (New York, N.Y.)*, 19(1), 36–50. <http://doi.org/10.1261/rna.035469.112>
- Mignone, F., Gissi, C., Liuni, S., & Pesole, G. (2002). Untranslated regions of mRNAs. *Genome Biology*, 3(3), REVIEWS0004. <http://doi.org/10.1186/gb-2002-3-3-reviews0004>
- Monani, U. R., Coover, D. D., & Burghes, A. H. (2000). Animal models of spinal muscular atrophy. *Human Molecular Genetics*, 9(16), 2451–2457. <http://doi.org/10.1177/0883073807305667>
- Monani, U. R., Pastore, M. T., Gavrilina, T. O., Jablonka, S., Le, T. T., Andreassi, C., ... Burghes, A. H. M. (2003). A transgene carrying an A2G missense mutation in the SMN gene modulates phenotypic severity in mice with severe (type I) spinal muscular atrophy. *Journal of Cell Biology*, 160(1), 41–52. <http://doi.org/10.1083/jcb.200208079>
- Monani, U. R., Sendtner, M., Coover, D. D., Parsons, D. W., Andreassi, C., Le, T. T., ... Burghes, A. H. M. (2000). The human centromeric survival motor neuron gene (SMN2) rescues embryonic lethality in Smn^{-/-} mice and results in a mouse with spinal muscular atrophy. *Human Molecular Genetics*, 9(3), 333–339. <http://doi.org/10.1093/hmg/ddm236>
- Munsat, T. L., & Davies, K. E. (1992). International SMA Consortium Meeting (26–28 June 1992, Bonn, Germany). *Neuromuscular Disorders*, 2(5–6), 423–428. [http://doi.org/10.1016/S0960-8966\(06\)80015-5](http://doi.org/10.1016/S0960-8966(06)80015-5)
- Ni, J. Q., Markstein, M., Binari, R., Pfeiffer, B., Liu, L. P., Villalta, C., ... Perrimon, N. (2008). Vector and parameters for targeted transgenic RNA interference in *Drosophila melanogaster*. *Nature Methods*, 5(1), 49–51. <http://doi.org/10.1038/nmeth1146>
- Nichols, C. D. (2006). *Drosophila melanogaster* neurobiology, neuropharmacology, and how the fly can inform central nervous system drug discovery, 112, 677–700. <http://doi.org/10.1016/j.pharmthera.2006.05.012>
- Nicholson, L., Singh, G. K., Osterwalder, T., Roman, G. W., Davis, R. L., & Keshishian, H. (2008). Spatial and temporal control of gene expression in *Drosophila* using the inducible geneSwitch GAL4 system. I. Screen for larval nervous system drivers. *Genetics*, 178(1), 215–234. <http://doi.org/10.1534/genetics.107.081968>
- Nilsson, J., Sengupta, J., Frank, J., & Nissen, P. (2004). Regulation of eukaryotic translation by the RACK1 protein: A platform for signalling molecules on the

- ribosome. *EMBO Reports*, 5(12), 1137–1141.
<http://doi.org/10.1038/sj.embor.7400291>
- Nussbacher, J. K., Batra, R., Lagier-tourenne, C., Yeo, G. W., Medicine, M., Diego, S., ... Jolla, L. (2015). RNA binding proteins in neurodegeneration: Seq and you shall receive. *Trends Neuroscience*, 38(4), 226–236.
<http://doi.org/10.1016/j.tins.2015.02.003.RNA>
- Okray, Z., & Hassan, B. A. (2013). Genetic approaches in *Drosophila* for the study neurodevelopmental disorders. *Neuropharmacology*, 68, 150–156.
<http://doi.org/10.1016/j.neuropharm.2012.09.007>
- Olesnicky, E. C., Killian, D. J., Garcia, E., Morton, M. C., Rathjen, A. R., Sola, I. E., & Gavis, E. R. (2014). Extensive Use of RNA-Binding Proteins in *Drosophila* Sensory Neuron Dendrite Morphogenesis. *G3: Genes, Genome, Genetics*, 4(2), 297–306. <http://doi.org/10.1534/g3.113.009795>
- Oliveira, M. (2010). *Caracterização bioquímica e funcional de complexos proteicos associados ao 3'UTR de mRNAs alvo*. Faculdade de Ciências da Universidade de Lisboa.
- Osterwalder, T., Yoon, K. S., White, B. H., & Keshishian, H. (2001). A conditional tissue-specific transgene expression system using inducible GAL4. *Proceedings of the National Academy of Sciences*, 98(22), 12596–12601.
<http://doi.org/10.1073/pnas.221303298>
- Park, G. H., Kariya, S., & Monani, U. R. (2010). Spinal muscular atrophy: new and emerging insights from model mice. *Current Neurology and Neuroscience Reports*, 10(2), 108–117. <http://doi.org/10.1007/s11910-010-0095-5.Spinal>
- Patrizi, A. L., Tiziano, F., Zappata, S., Donati, M. A., Neri, G., & Brahe, C. (1999). SMN protein analysis in fibroblast, amniocyte and CVS cultures from spinal muscular atrophy patients and its relevance for diagnosis. *European Journal of Human Genetics*, 7(3), 301–309. <http://doi.org/10.1038/sj.ejhg.5200286>
- Paushkin, S., Gubitz, A. K., Massenet, S., & Dreyfuss, G. (2002). The SMN complex, an assemblyosome of ribonucleoproteins. *Current Opinion in Cell Biology*, 14(3), 305–312. [http://doi.org/10.1016/S0955-0674\(02\)00332-0](http://doi.org/10.1016/S0955-0674(02)00332-0)
- Pellizzoni, L. (2007). Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Reports*, 8(4), 340–345.
<http://doi.org/10.1038/sj.embor.7400941>
- Pellizzoni, L., Kataoka, N., Charroux, B., & Dreyfuss, G. (1998). A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell*, 95(5), 615–624. [http://doi.org/10.1016/S0092-8674\(00\)81632-3](http://doi.org/10.1016/S0092-8674(00)81632-3)
- Perrimon, N., Ni, J., Perkins, L., Noller, H. F., Volpe, T., & Martienssen, R. A. (2013). In vivo RNAi : Today and Tomorrow In vivo RNAi : Today and Tomorrow, 1–11. <http://doi.org/10.1101/cshperspect.a003640>
- Phelps, C. B., & Brand, A. H. (1998). Ectopic Gene Expression in *Drosophila* Using GAL4 System. *Methods*, 14(4), 367–379.
<http://doi.org/10.1006/meth.1998.0592>
- Potter, C. J., Tasic, B., Russler, E. V., Liang, L., & Luo, L. (2010). The Q System: A Repressible Binary System for Transgene Expression, Lineage Tracing and Mosaic Analysis. *Cell*, 141(3), 536–548.
<http://doi.org/10.1016/j.cell.2010.02.025.The>
- Press, T. (1975). Pathogenesis of Motor Neurone Disease Involuntary Facial Movements Narcolepsy, (March), 475–476.

- Prüßing, K., Voigt, A., & Schulz, J. B. (2013). *Drosophila melanogaster* as a model organism for Alzheimer's disease. *Molecular Neurodegeneration*, 8(1). <http://doi.org/10.1186/1750-1326-8-35>
- Rajendra, T. K., Gonsalvez, G. B., Walker, M. P., Shpargel, K. B., Salz, H. K., & Matera, A. G. (2007). A *Drosophila melanogaster* model of spinal muscular atrophy reveals a function for SMN in striated muscle. *Journal of Cell Biology*, 176(6), 831–841. <http://doi.org/10.1083/jcb.200610053>
- Ramanathan, A., Robb, G. B., & Chan, S. H. (2016). mRNA capping: Biological functions and applications. *Nucleic Acids Research*, 44(16), 7511–7526. <http://doi.org/10.1093/nar/gkw551>
- Raška, I., Andrade, L. E. C., Ochs, R. L., Chan, E. K. L., Chang, C. M., Roos, G., & Tan, E. M. (1991). Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Experimental Cell Research*, 195(1), 27–37. [http://doi.org/10.1016/0014-4827\(91\)90496-H](http://doi.org/10.1016/0014-4827(91)90496-H)
- Ready, D. F., Hanson, T. E., & Benzer, S. (1976). Development of the *Drosophila* retina, a neurocrystalline lattice. *Developmental Biology*, 53(2), 217–240. [http://doi.org/10.1016/0012-1606\(76\)90225-6](http://doi.org/10.1016/0012-1606(76)90225-6)
- Rochette, C. F., Gilbert, N., & Simard, L. R. (2001). SMN gene duplication and the emergence of the SMN2 gene occurred in distinct hominids: SMN2 is unique to *Homo sapiens*. *Human Genetics*, 108(3), 255–266. <http://doi.org/10.1007/s004390100473>
- Roman, G., Endo, K., Zong, L., & Davis, R. L. (2001). P[Switch], a system for spatial and temporal control of gene expression in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, 98(22), 12602–7. <http://doi.org/10.1073/pnas.221303998>
- Rossoll, W., Jablonka, S., Andreassi, C., Kröning, A. K., Karle, K., Monani, U. R., & Sendtner, M. (2003). Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of β -actin mRNA in growth cones of motoneurons. *Journal of Cell Biology*, 163(4), 801–812. <http://doi.org/10.1083/jcb.200304128>
- Rubin, G. M., & Lewis, E. B. (2000). A Brief History of *Drosophila* 's Contributions to Genome Research, 287(March), 2216–2219.
- Russman, B. S. (2007). Spinal muscular atrophy: clinical classification and disease heterogeneity. *Journal of Child Neurology*, 22(8), 946–951. <http://doi.org/10.1177/0883073807305673>
- Sainsbury, S., Bernecky, C., & Cramer, P. (2015). Structural basis of transcription initiation by RNA polymerase II. *Nature Reviews Molecular Cell Biology*, 16(3), 129–143. <http://doi.org/10.1038/nrm3952>
- Santa-Cruz, L. D., & Tapia, R. (2014). Role of Energy Metabolic Deficits and Oxidative Stress in Excitotoxic Spinal Motor Neuron Degeneration in Vivo. *ASN Neuro*, 6(2), 83–93. <http://doi.org/10.1042/AN20130046>
- Schäffler, K., Schulz, K., Hirmer, A., Wiesner, J., Grimm, M., Sickmann, A., & Fischer, U. (2010). A stimulatory role for the La-related protein 4B in translation. *RNA (New York, N.Y.)*, 16(8), 1488–99. <http://doi.org/10.1261/rna.2146910>
- Schrank, B., Götz, R., Gunnensen, J. M., Ure, J. M., Toyka, K. V., Smith, A. G., & Sendtner, M. (1997). Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proceedings of the National Academy of Sciences of the United States of America*, 94(18), 9920–5.

- <http://doi.org/10.1073/pnas.94.18.9920>
- Sen, A., Dimlich, D. N., Guruharsha, K. G., Kankel, M. W., Hori, K., Yokokura, T., ... Artavanis-Tsakonas, S. (2013). Genetic circuitry of Survival motor neuron, the gene underlying spinal muscular atrophy. *Proceedings of the National Academy of Sciences*, 110(26), E2371–E2380. <http://doi.org/10.1073/pnas.1301738110>
- Sen, A., Yokokura, T., Kankel, M. W., Dimlich, D. N., Manent, J., Sanyal, S., & Artavanis-Tsakonas, S. (2011). Modeling spinal muscular atrophy in *Drosophila* links Smn to FGF signaling. *Journal of Cell Biology*, 192(3), 481–495. <http://doi.org/10.1083/jcb.201004016>
- Sendtner, M. (2005). Motor Neuron Diseases: Molecular Mechanism , Pathophysiology , and Treatments Keywords. *Review in Cell Biology and Molecular Medicine*, 559–574. <http://doi.org/10.1002/3527600906>
- Shababi, M., Lorson, C. L., & Rudnik-Schöneborn, S. S. (2014). Spinal muscular atrophy: A motor neuron disorder or a multi-organ disease? *Journal of Anatomy*, 224(1), 15–28. <http://doi.org/10.1111/joa.12083>
- Shaw, P. (2005). Molecular and cellular pathways of neurodegeneration in motor neurone disease. *Journal of Neurology, Neurosurgery and Psychiatry*, 76(8), 1046–1057. <http://doi.org/10.1136/jnnp.2004.048652>
- Shaw, P. J. (1999). Motor neurone disease. *BMJ (Clinical Research Ed.)*, 318(7191), 1118–21. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1115517&tool=pmcentrez&rendertype=abstract>
- Shaw, P. J., & Eggett, C. J. (2000). Molecular factors underlying selective vulnerability of motor neurons to neurodegeneration in amyotrophic lateral sclerosis. *Journal of Neurology*, 247 Suppl, 117–127. <http://doi.org/10.1007/BF03161151>
- Squire, L. R., Bloom, F. E., Spitzer, N. C., Lac, S. du, Ghosh, A., & Berg, D. (2008). *Fundamental Neuroscience* (3^a). Elsevier Inc. <http://doi.org/10.1016/B978-0-12-385870-2.00032-9>
- St Johnston, D. (2002). The art and design of genetic screens: *Drosophila melanogaster*. *Nature Reviews Genetics*, 3(3), 176–188. <http://doi.org/10.1038/nrg751>
- Statland, J. M., Barohn, R. J., McVey, A. L., Katz, J. S., & Dimachkie, M. M. (2015). Patterns of Weakness, Classification of Motor Neuron Disease, and Clinical Diagnosis of Sporadic Amyotrophic Lateral Sclerosis. *Neurologic Clinics*, 33(4), 735–748. <http://doi.org/10.1016/j.ncl.2015.07.006>
- Stavraka, C., & Blagden, S. (2015). The La-related proteins, a family with connections to cancer. *Biomolecules*, 5(4), 2701–2722. <http://doi.org/10.3390/biom5042701>
- Stifani, N. (2014). Motor neurons and the generation of spinal motor neuron diversity. *Frontiers in Cellular Neuroscience*, 8(293), 1–22. <http://doi.org/10.3389/fncel.2014.00293>
- Sugarman, E. A., Nagan, N., Zhu, H., Akmaev, V. R., Zhou, Z., Rohlf, E. M., ... Allitto, B. A. (2011). Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: clinical laboratory analysis of 400 specimens. *European Journal of Human Genetics*, 20(10), 27–32. <http://doi.org/10.1038/ejhg.2011.134>
- Sumner, C. J. (2007). Molecular Mechanisms of Spinal Muscular Atrophy. *Journal*

- of Child Neurology*, 22(8), 979–989. <http://doi.org/10.1083/jcb.200703187>
- Sumner, C., Paushkin, S., & Ko, C.-P. (2017). *Spinal muscular atrophy* (1^o). Elsevier Inc.
- Szűts, D., & Bienz, M. (2000). LexA chimeras reveal the function of Drosophila Fos as a context-dependent transcriptional activator. *Proceedings of the National Academy of Sciences of the United States of America*, 97(10), 5351–6. <http://doi.org/10.1073/pnas.97.10.5351>
- Tan, F. H. P., & Azzam, G. (2017). Drosophila melanogaster: Deciphering Alzheimer's disease. *Malaysian Journal of Medical Sciences*, 24(2), 6–20. <http://doi.org/10.21315/mjms2017.24.2.2>
- Taylor, J. E., Thomas, N. H., Lewis, C. M., Abbs, S. J., Rodrigues, N. R., Davies, K. E., & Mathew, C. G. (1998). Correlation of SMNt and SMNc gene copy number with age of onset and survival in spinal muscular atrophy. *European Journal of Human Genetics: EJHG*, 6(5), 467–474. <http://doi.org/10.1038/sj.ejhg.5200210>
- Tiziano, F. D., Melki, J., & Simard, L. R. (2013). Solving the puzzle of spinal muscular atrophy: What are the missing pieces? *American Journal of Medical Genetics, Part A*, 161(11), 2836–2845. <http://doi.org/10.1002/ajmg.a.36251>
- Turner, M., & Díaz-Muñoz, M. D. (2018). RNA-binding proteins control gene expression and cell fate in the immune system. *Nature Immunology*, 19(2), 120–129. <http://doi.org/10.1038/s41590-017-0028-4>
- Valetdinova, K. R., Medvedev, S. P., & Zakian, S. M. (2015). Model systems of motor neuron diseases as a platform for studying pathogenic mechanisms and searching for therapeutic agents. *Acta Naturae*, 7(1), 19–36.
- Venken, K. J. T., Simpson, J. H., & Bellen, H. J. (2011). Genetic manipulation of genes and cells in the nervous system of the fruit fly. *Neuron*, 72(2), 202–230. <http://doi.org/10.1016/j.neuron.2011.09.021>
- Watihayati, M., Fatemeh, H., Marini, M., Atif, A. B., Zahiruddin, W. M., Sasongko, T. H., ... Zilfalil, B. A. (2009). Combination of SMN2 copy number and NAIP deletion predicts disease severity in spinal muscular atrophy. *Brain and Development*, 31(1), 42–45. <http://doi.org/10.1016/j.braindev.2008.08.012>
- Will, C. L., & Lührmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. *Current Opinion in Cell Biology*, 13(3), 290–301. [http://doi.org/10.1016/S0955-0674\(00\)00211-8](http://doi.org/10.1016/S0955-0674(00)00211-8)
- Worms, P. M. (2001). The epidemiology of motor neuron diseases: a review of recent studies. *Journal of the Neurological Sciences*, 191(1–2), 3–9. [http://doi.org/10.1016/S0022-510X\(01\)00630-X](http://doi.org/10.1016/S0022-510X(01)00630-X)
- Yang, R., Gaidamakov, S. A., Xie, J., Lee, J., Martino, L., Kozlov, G., ... Maraia, R. J. (2011). La-Related Protein 4 Binds Poly(A), Interacts with the Poly(A)-Binding Protein MLLE Domain via a Variant PAM2w Motif, and Can Promote mRNA Stability. *Molecular and Cellular Biology*, 31(3), 542–556. <http://doi.org/10.1128/MCB.01162-10>
- Yong, J., Golembe, T. J., Battle, D. J., Pellizzoni, L., & Dreyfuss, G. (2004). snRNAs contain specific SMN-binding domains that are essential for snRNP assembly. *Molecular and Cellular Biology*, 24(7), 2747–2756. <http://doi.org/10.1128/MCB.24.7.2747-2756.2004>
- Zhang, H. L., Pan, F., Hong, D., Shenoy, S. M., Singer, R. H., & Bassell, G. J. (2003). Active transport of the survival motor neuron protein and the role of exon-7 in cytoplasmic localization. *The Journal of Neuroscience: The Official Journal*

of the Society for Neuroscience, 23(16), 6627–37. <http://doi.org/23/16/6627>
[pii]
Zhou, Q., Li, T., Price, & David. (2012). RNA Polymerase II Elongation Control. *Annual Review Biochemistry*, 81, 119–143. <http://doi.org/10.1146/annurev-biochem-052610-095910>.

VIII. **Anexo**

Anexo I: Sequência do vetor de expressão pUAST_CG11505

Sequência

AAGCTTGGGCTGCATCTCTCCGGATCCAAGCTTGCATGCCTGCAGGTCGGAGTACTGT
CCTCCGAGCGGAGTACTGTCTCCGAGCGGAGTACTGTCTCCGAGCGGAGTACTGTC
CTCCGAGCGGAGTACTGTCTCCGAGCGGAGACTCTAGCGAGCGCCGGAGTATAAAT
AGAGGCGCTTCGTCTACGGAGCGACAATTCAATTCAAACAAGCAAAGTGAACACGTC
GCTAAGCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAACAATCTGCA
GTAAAGTGCAAGTTAAAGTGAATCAATTAAAAGTAACCAGCAACCAAGTAAATCAAC
TGCAACTACTGAAATCTGCCAAGAAGTAATTATTGAATACAAGAAGAGAACTCTGAA
TAGGGAATTGGGAATTCGTTAACAGATCTTGCGGCCCGC**ATGTACCCATACGATGTTCC**
AGATTACGCTATGAACGGAGATGCACTCAAGATCCAGCCGCCGTCTACACCAACGTTT
CTGTCGAGACGGCGATGCTAGCAACTAATCTGTTTGGCGCCACAACCCAAATTGCAGC
CTCACCCGCAGCAGCCGCCGTGCCACATACCGTTGATCACAGCCGCAGCTGCAGCC
GCCGCCGCCGCAGGCGCTCCAGCAGTAGCTCCTCCAGTTGCCGTTGCTCCAGTGGCCA
CCGCATCAGCAGCACAAGCAGCAGTTGTGCCCCGCCAACAACAGCAGCAGGCGCATCC
ACACCAGGCCCAGATTTTGGCCCATACGCACGTAGCCACCAACAACAACAGCAGCAT
CAGCAGCAGCAGCAGCAGACGATCCAACAGCATCTCCACCAGCAGCAGCAACAACAG
TCGCCGCATCCCGCCCAGCATCTGACCTACGGCCATCAGCCGGCGCTTTCGCAGGCCA
CGGCTGGAGGATCAGGAACAATTTCCGGAGGAGCAGTAGGACCATCCGGAGTAGTCG
TGGAGGGCCAGGACACCAACGAGTACGCCATCATGAATGGCGCACTCGGCCAGACCC
AGGATGGCACTGTGGTCTGCTACACCACAGACGCCCTGAACAACACGAATTTAGTTG
CACTCGATCCGCAGCCCCATCAAATTGTAGTGCCGGTCCAGGTTCTGTCCAAGTCCC
AGTGCCCGTTTCAGCTACCAGCAAATGGATCCGCTGATCCGCAGCAAGGCAGCCACAAT
GCAGCTGGTGGAGAGGAGCCCAACATACCGCTGGACAAGCTCAAACAATGCTGGCC
ACCCAACTTGAGTACTACTTCTCCAGGGAGAACCTGGCAAACGACACCTACCTGCTCT
CCCAGATGGACAGCGACCAGTATGTGCCATTTACACTGTGGCCAGATTTAATTTGG
TGCGAAAGCTAACCAATGACATCAATCTCATCACTGAGGTGCTGCGTGAATCGCCCA
ATGTCCAGGTGGACGACAAAGGCCTGCGCGTGCGTCCCAATCGCAAGCGTTGCATCAT
CATTCTGCGCGAAATATCGAACAACACGCCTCTCGATGACGTCAAGGCTCTGTTCAGC
AATGAGAGCTGCCCCCGCCGATATCCTGCGAATTTCGCCGCCAATAACTCGTGGTACA
TAACCTTCGAGTCGGACGAAGATGCGCAAAGGCGTATAAGTACCTGCGCGAAGAGG

TCAAGGAGTTCCAGGGCAAGCCGATCATGGCTCGCATCAAGCCCAAGTCGTTTATTA
ATCGCATCCAAGCTGTACCAAAGAACGGCTATCGCCTCACGTCGCCGCCGACTGCTAA
TGTCTTTGATCCAGCCGGAGCCGGTGGCGCCACAGTTAGCTATGCGGCTGCCCAGCAG
CCCCGCTATCTGTACACCAACGGAGCGGCCATCGCGGCTCCCGCTTCAGTTCAGTACA
GCAACCCAGTACTGATTCCGATTCCGCAGAATCAATTCTACCCTGGATTGGTGGCTGG
TCCCTGGCCGCCTGGCACTGTAGCAGCTACTACCGCCCATGGGCAAACTTCTACGAG
CTCGGAGGAAACATATTTACCACCAATCCTTTACCTCCGCAGCCAGTGACAGCGGGAT
TCGCTCAAGCTCCACCACCCACCTCACAAGCAATGGTCACGCCCAAACCGCAGGGCGG
TGGGCGCTACAATAACAATCACCGCGGCAATCCAAACAACGTTGGCGGTGCAAACCTC
GGGACCACGCGCCGAGTCGAGGAGCAAACCACCGCGCAGCACACCATCAGCCTCCCAT
ATCCAGGGCGGTAACATCATGCCATCCAGATAACTGCACCCATTCCAGGAGGAATG
GTTAGCGTGGTTCCGACCAACATTGTGCAGGATCCATTACAGCCGGCACCTTTGCAAC
AGCAACCCCAGCAGGTGATCCAACAAATTCAGCAGCAGCCATCGAGCTCAGCTCAGCA
GCAACAGGCCCGCCACGGTGCAGATGAACAACACTACGCGTCACTACCCCATCAAGAGC
AACTGGAAGGGTGGCATGCAAAAGAACCTAGACAAGAGCTATAGCGGTGGTGTGGCC
ACCCATCATCACTACACCACCCAGGTGCAGGCTCTGCCCCGCCACAGCCATCATCTAC
AGGCTCAGCAGCAGCTGCAGGGACAGACCCAGCAGCAGCACTATCAATTGGCTTCCTC
CAGCGCTGCAAGCGCTCCGCAAGTTACCTATGTAACCACTGCACCGGCCCCACAGCAG
CCGGGCCAGCATCAACACCACTTGGTGGTCCAGCAAACGCAGCCACAGCAGCAGCAAC
AACAGCAACAGGTAGTCCTTCAGCAGCAGCAAGTGCAGGTGCAGGAGCTGCAGGAAG
CTGGGGACGGCGTCGAGCATAGCTCTCACGCCATGCAGCAGGTGAATCGCAGTAGCA
ACATGTCTGCCAGCTCCTCCTCATCGCTGGCCACTTCGATGAGCAAGGAGCCACTGCA
ATGGCAGAATCGTCCGCGTCGTCGTCGTCGTCGATGAGGAGGGCGGTATCAACTATTC
CCCAGGCGGGCGAGTTGGGCTCTACGGCAGCTCGCCCTCAAACCCGCATCCGCAGCAG
CATCTGATGTCATCTTCCACCGGGAGCAATGTACAATCAGCTGGCGGAACCGATGGC
GGCGGAGCCTCCCATCGCGGCGGAGAGCGCCAGAGCCACTACAACACCATCCACGACG
TCCCACCGCCCCAGCACCGTGGCAACTACAAGGGCAATAACTACAATCCGCATTATCA
TGCCCAGCATTCTTCGATGGCGAGTGGCTCGCACCATCACCATCACCACAACGCCCTG
TCGTCGCAGCAGCAGCAACATCACTTGACCACAGGTGGCACAGC**ACCACAAGGCCGC**
GGCCTCTTCTGATCTAGAGGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACA
TAATTGGACAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTAA
AGTGTATAATGTGTTAACTACTGATTCTAATTGTTTGTGTATTTTAGATTCCAACC
TATGGAACCTGATGAATGGGAGCAGTGGTGGAATGCCTTTAATGAGGAAAACCTGTTT

TGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCT
ACTCCTCCAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAGAAATTG
CTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATT
TACACCACAAAGGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTTG
ATGTATAGTGCCTTGACTAGAGATCATAATCAGCCATACCACATTTGTAGAGGTTTT
ACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCA
ATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGC
ATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTTGTCC
AAACTCATCAATGTNNCTTATCATGTCTGGATCCACTAGAAGGCCTAATTC AATTCT
AGTATGTATGTAAGTTAATAAAACCCTTTTTTGGAGAATGTAGATTTAAAAAACAT
ATTTTTTTTTTTATTTTTTACTGCACTGGATATCATTGAACTTATCTGATCAGTTTTA
AATTTACTTCGATCCAAGGGTATTTGAAGTACCAGGTTCTTTCGATTACCTCTCACTC
AAAATGACATTCCACTCAAAGTCAGCGCTGTTTGCCTCCTTCTCTGTCCACAGAAATA
TCGCCGTCTCTTTCGCCGCTGCGTCCGCTATCTCTTTCGCCACCGTTTGTAGCGTTAC
CTAGCGTCAATGTCCGCCTTCAGTTGCACTTTGTGAGCGGTTTCGTGACGAAGCTCCA
AGCGGTTTACGCCATCAATTAAACACAAAGTGCTGTGCCAAAACCTCTCGCTTCTT
ATTTTTGTTTGTTTTTTTGAGTGATTGGGGTGGTGATTGGTTTTGGGTGGGTAAGCAG
GGGAAAGTGTA AAAATCCCGGCAATGGGCCAAGAGGATCAGGAGCTATTAATTTCG
GGAGGCAGCAAACACCCATCTGCCGAGCATCTGAACAATGTGAGTAGTACATGTGCA
TACATCTTAAGTTCACTTGATCTATAGGAACTGCGATTGCAACATCAAATTGTCTGC
GGCGTGAGAACTGCGACCCACAAAAATCCCAAACCGCAATCGCACAAACAAATAGTG
ACACGAAACAGATTATTCTGGTAGCTGTGCTCGCTATATAAGACAATTTTTTAAGATC
ATATCATGATCAAGACATCTAAAGGCATTCATTTTCGACTACATTCTTTTTTACAAA
AAATATAACAACCAGATATTTTAAGCTGATCCTAGATGCACAAAAAATAAATAAAAG
TATAAACCTACTTCGTAGGATACTTCGTTTTGTTCGGGGTTAGATGAGCATAACGCT
TG TAGTTGATATTTGAGATCCCCTATCATTGCAGGGTGACAGCGGACGCTTCGCAGA
GCTGCATTAAACCAGGGCTTCGGGCAGGCCAAAAACTACGGCACGCTCCTGCCACCCAG
TCCGCCGGAGGACTCCGGTTCAGGGAGCGGCCAACTAGCCGAGAACCTCACCTATGCC
TGGCACAATATGGACATCTTTGGGGCGGTCAATCAGCCGGGCTCCGGATGGCGGCAG
CTGGTCAACCGGACACGCGGACTATTCTGCAACGAGCGACACATACGGGCGCCAGGA
AACATTTGCTCAAGAACGGTGAGTTTCTATTTCGAGTCGGCTGATCTGTGTGAAATC
TTAATAAAGGGTCCAATTACCAATTTGAAACTCAGTTTGCGGCGTGGCCTATCCGGG
CGAACTTTTGGCCGTGATGGGCAGTTCCGGTGCCGAAAGACGACCCTGCTGAATGCC

CTTGCCCTTTCGATCGCCGCAGGGCATCCAAGTATCGCCATCCGGGATGCGACTGCTCA
ATGGCCAACCTGTGGACGCCAAGGAGATGCAGGCCAGGTGCGCCTATGTCCAGCAGG
ATGACCTCTTTATCGGCTCCCTAACGGCCAGGGAACACCTGATTTTCCAGGCCATGGT
GCGGATGCCACGACATCTGACCTATCGGCAGCGAGTGGCCCCGCTGGATCAGGTGATC
CAGGAGCTTTCGCTCAGCAAATGTCAGCACACGATCATCGGTGTGCCCGGCAGGGTG
AAAGGTCTGTCCGGCGGAGAAAGGAAGCGTCTGGCATTTCGCCTCCGAGGCACTAACC
GATCCGCCGCTTCTGATCTGCGATGAGCCACCTCCGGACTGGACTCATTTACCGCCC
ACAGCGTCGTCCAGGTGCTGAAGAAGCTGTTCGAGAAGGGCAAGACCGTCATCCTGA
CCATTCATCAGCCGTCTTCCGAGCTGTTTGAGCTCTTTGACAAGATCCTTCTGATGGC
CGAGGGCAGGGTAGCTTTCTTGGGCACTCCCAGCGAAGCCGTCGACTTCTTTTCCTAG
TGAGTTCGATGTGTTTATTAAGGGTATCTAGCATTACATTACATCTCAACTCCTATC
CAGCGTGGGTGCCAGTGTCCTACCAACTACAATCCGGCGGACTTTTACGTACAGGTG
TTGGCCGTGTGCCCCGACGGGAGATCGAGTCCCGTGATCGGATCGCCAAGATATGG
CGACAATTTTGCTATTAGCAAAGTAGCCCGGGATATGGAGCAGTTGTTGGCCACCAA
AAATTTGGAGAAGCCACTGGAGCAGCCGGAGAATGGGTACACCTACAAGGCCACCTG
GTTTCATGCAGTTCCGGGCGGTCTGTGGCGATCCTGGCTGTTCGGTGCTCAAGGAACCA
CTCCTCGTAAAAGTGCGACTTATTCAGACAACGGTGAGTGGTTCCAGTGGAACAAA
TGATATAACGCTTACAATTCTTGGAACAAATTTCGCTAGATTTTAGTTAGAATTGCC
TGATTCCACACCTTCTTAGTTTTTTTCAATGAGATGTATAGTTTATAGTTTTGCAG
AAAATAAATAAATTTCAATTAACCTCGCGAACATGTTGAAGATATGAATATTAATGAG
ATGCGAGTAACATTTTAATTTGCAGATGGTTGCCATCTTGATTGGCCTCATCTTTTT
GGGCAACAACCTCACGCAAGTGGGCGTGATGAATATCAACGGAGCCATCTTCCTCTTC
CTGACCAACATGACCTTTCAAACGTCTTTGCCACGATAAATGTAAGTCTTGTTTAG
AATACATTTGCATATTAATAATTTACTAACTTTCTAATGAATCGATTGATTTAGGT
GTTACCTCAGAGCTGCCAGTTTTTATGAGGGAGGCCCCGAAGTCGACTTTATCGCTGT
GACACATACTTTCTGGGCAAAACGATTGCCGAATTACCGCTTTTTCTCACAGTGCCAC
TGGTCTTCACGGCGATTGCCTATCCGATGATCGGACTGCGGGCCGGAGTGCTGCACTT
CTTCAACTGCCTGGCGCTGGTCACTCTGGTGGCCAATGTGTCAACGTCCTTCGGATAT
CTAATATCCTGCGCCAGCTCCTCGACCTCGATGGCGCTGTCTGTGGGTCCGCCGGTTA
TCATACCATTCCTGCTCTTTGGCGGCTTCTTCTTGAACTCGGGCTCGGTGCCAGTATA
CCTCAAATGGTTGTTCGTACCTCTCATGGTTCCGTTACGCCAACGAGGGTCTGCTGATT
AACCAATGGGCGGACGTGGAGCCGGGCGAAATTAGCTGCACATCGTCGAACACCACG
TGCCCCAGTTCGGGCAAGGTCATCCTGGAGACGCTTAACTTCTCCGCCGCCGATCTGC

CGCTGGACTACGTGGGTCTGGCCATTCTCATCGTGAGCTTCCGGGTGCTCGCATATCT
GGCTCTAAGACTTCGGGCCCCGACGCAAGGAGTAGCCGACATATATCCGAAATAACTG
CTTGTTTTTTTTTTTACCATTATTACCATCGTGTTTACTGTTTATTGCCCCCTCAAAA
AGCTAATGTAATTATATTTGTGCCAATAAAAAACAAGATATGACCTATAGAATACAAG
TATTTCCCCTTCGAACATCCCCACAAGTAGACTTTGGATTTGTCTTCTAACCAAAAGA
CTTACACACCTGCATACCTTACATCAAAAACCTCGTTTATCGCTACATAAAACACCGGG
ATATATTTTTTATATACATACTTTTCAAATCGCGCGCCCTCTTCATAATTCACCTCCA
CCACACCACGTTTCGTAGTTGCTCTTTCGCTGTCTCCACCCGCTCTCCGCAACACAT
TCACCTTTTGTTTCGACGACCTTGGAGCGACTGTCGTTAGTTCCGCGCGATTCCGGTTCG
CTCAAATGGTTCCGAGTGGTTCATTTTCGTCTCAATAGAAATTAGTAATAAATATTTG
TATGTACAATTTATTTGCTCCAATATATTTGTATATATTTCCCTCACAGCTATATTT
ATTCTAATTTAATATTATGACTTTTTTAAGGTAATTTTTTGTGACCTGTTCCGGAGTGA
TTAGCGTTACAATTTGAACTGAAAGTGACATCCAGTGTTTGTTCCTTGTGTAGATGC
ATCTCAAAAAAATGGTGGGCATAATAGTGTTGTTTATATATATCAAAAAATAACAACT
ATAATAATAAGAATACATTTAATTTAGAAAATGCTTGGATTTCACTGGAAGTAGAAT
TAATTCGGCTGCTGCTCTAAACGACGCATTTTCGTACTCCAAAGTACGAATTTTTTCCC
TCAAGCTCTTATTTTCATTAAACAATGAACAGGACCTAACGCACAGTCACGTTATTG
TTTACATAAATGATTTTTTTTTTACTATTCAAACCTTACTCTGTTTGTGTACTCCCACTGG
TATAGCCTTCTTTTATCTTTTTCTGGTTCAGGCTCTATCACTTTACTAGGTACGGCATC
TGCGTTGAGTCGCCTCCTTTTAAATGTCTGACCTTTTGCAGGTGCAGCCTTCCACTGC
GAATCTTTAAAGTGGGTATCACAAATTTGGGAGTTTTTCACCAAGGCTGCACCCAAGG
CTCTGCTCCCACAATTTTCTCTTAATAGCACACTTCGGCACGTGAATTAATTTTACTC
CAGTCACAGCTTTGCAGCAAAATTTGCAATATTTCAATTTTTTTTTTATTCCACGTAAG
GGTTAATGTTTTCAAAAAAAAATTCGTCCGCACACAACCTTTCCTCTCAACAAGCAA
ACGTGCACTGAATTTAAGTGTATACTTCGGTAAGCTTCGGCTATCGACGGGACCACC
TTATGTTATTTTCATCATGGGCCAGACCCACGTAGTCCAGCGGCAGATCGGCGGCGGA
GAAGTTAAGCGTCTCCAGGATGACCTTGCCCGAACTGGGGCACGTGGTGTTCGACGA
TGTGCAGCTAATTTGCCCCGGCTCCACGTCCGCCCATTGGTTAATCAGCAGACCCTCG
TTGGCGTAACGGAACCATGAGAGGTACGACAACCATTTGAGGTATACTGGCACCGAG
CCCGAGTTCAAGAAGAAGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACA
AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG
CGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA
TACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTA

GGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
CGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTA
AGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGA
AGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTT
GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGC
AAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCT
ACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGA
TTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA
ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAG
GCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCTG
TGTAATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAC
CGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAG
GGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGT
TGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCC
ATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTACAGTCCG
GTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTA
GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGAGTGTTATCACTCAT
GGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCT
GTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT
TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAA
GTGCTCATCATTGGAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG
TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTA
CTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGG
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATT
GAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGA
AAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCT
AAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCT
TTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGA
GACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGC
GTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGAT
TGTAAGTGTGAGAGTGCACCATATGCGGTGTGAAATACCGCACCGAATCGCGCGGAATA
ACGACAGTCGCTCCAAGGTCGTCTGAACAAAAGGTGAATGTGTTGCGGAGAGCGGGTG

GGAGACAGCGAAAGAGCAACTACGAAACGTGGTGTGGTGGAGGTGAATTATGAAGA
GGGCGCGCGATTTGAAAAGTATGTATATAAAAAATATATCCCGGTGTTTTATGTAGC
GATAAACGAGTTTTTGTATGTAAGGTATGCAGGTGTGTAAGTCTTTTGGTTAGAAGAC
AAATCCAAAGTCTACTTGTGGGGATGTTCTGAAGGGGAAATACTTGTATTCTATAGGT
CATATCTTGTTTTTATTGGCACAAATATAATTACATTAGCTTTTTGAGGGGGCAATA
AACAGTAAACACGATGGTAATAATGGTAAAAAAAAAAAAACAAGCAGTTATTTCTGGAT
ATATGTCGGCTACTCCTTGCGTCGGGCCCCGAAGTCTTAGAGCCAGATATGCGAGCACC
CGGAAGCTCACGATGAGAATGGCCAGACCATGATGAAATAACATAAGGTGGTCCCGT
CGGCAAGAGACATCCACTTAACGTATGCTTGCAATAAGTGCGAGTGAAAGGAATAGT
ATTCTGAGTGTCGTATTGAGTCTGAGTGAGACAGCGATATGATTGTTGATTAACCCT
TAGCATGTCCGTGGGGTTTGAATTAACTCATAATATTAATTAGACGAAATTATTTTT
AAAGTTTTATTTTAAATAATTTGCGAGTACGCA

Legenda:

Primer Clonagem Forward (com NotI + **HA tag**)

Sequência do CG11505

Primer Clonagem Reverse (com XbaI)